

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:43:59 ; Search time 11 Seconds
(without alignments)
1315.931 Million cell updates/sec

Title: US-10-027-059A-1

Perfect score: 1889

Sequence: 1 MADADGFLAHTPLEPDAK.....CDHPWSCFLSLFLSLGWSG 349

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1767	93.5	518	1	TXB5_HUMAN
2	1712	90.6	518	1	TXB5_MOUSE
3	1130.5	59.8	545	1	TXB4_HUMAN
4	865.5	45.8	173	1	TXB4_MOUSE
5	658.5	34.9	423	1	TXB2_CAEL
6	653.5	34.6	414	1	TXB3_CHICK
7	652.5	34.5	361	1	TXB1_CHICK
8	649.5	34.4	701	1	TXB2_MOUSE
9	635.5	33.6	436	1	TXB6_HUMAN
10	633.5	33.5	702	1	TXB2_HUMAN
11	629.5	33.3	540	1	TXB6_MOUSE
12	626	33.1	398	1	TXB1_HUMAN
13	626	33.1	988	1	OMB_DROME
14	624.5	33.1	742	1	TXB3_HUMAN
15	616.5	32.6	251	1	TX20_HUMAN
16	616.5	32.6	297	1	TX20_MOUSE
17	606	32.1	455	1	VEGT_XENLA
18	603.5	31.9	613	1	TX18_MOUSE
19	601	31.8	501	1	TX18_HUMAN
20	591.5	31.3	473	1	TXB6_BRARE
21	589	31.2	602	1	TX15_MOUSE
22	587.5	31.1	184	1	TXB1_MOUSE
23	583.5	30.9	181	1	TXB3_MOUSE
24	575	30.4	681	1	BRC2_HALRO
25	558	29.5	177	1	TX13_MOUSE
26	549	29.1	346	1	TX12_CAEL
27	544.5	28.8	436	1	BRAC_MOUSE
28	542.5	28.7	435	1	BRAC_HUMAN
29	541.5	28.7	432	1	BRAC_XENLA
30	541.5	28.7	544	1	H15_DROME
31	536.5	28.4	423	1	BRAC_BRARE
32	536.5	28.4	681	1	TBR1_MOUSE
33	536.5	28.4	682	1	TBR1_HUMAN
					Q95933 homo sapien
					P70326 mus musculus
					P57082 homo sapien
					P70325 mus musculus
					Q19691 caenorhabdi
					O73718 gallus gall
					P79779 gallus gall
					Q60707 mus musculus
					O95947 homo sapien
					Q13207 homo sapien
					P70327 mus musculus
					O43435 homo sapien
					Q24432 drosophila
					O15119 homo sapien
					Q9umr3 homo sapien
					Q9es03 mus musculus
					P87377 xenopus lae
					Q9ep26 mus musculus
					O95935 homo sapien
					P79742 brachydanio
					O70306 mus musculus
					P70323 mus musculus
					P70324 mus musculus
					O01409 halocynthia
					O54841 mus musculus
					P90971 caenorhabdi
					P20293 mus musculus
					O15178 homo sapien
					P24781 xenopus lae
					Q94890 drosophila
					Q07998 brachydanio
					Q64336 mus musculus
					Q16650 homo sapien

RESULT 1

ID	TXB5_HUMAN	STANDARD:	PRT:	518 AA.
AC	Q95933; Q9Y412; O15301;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	T-box transcription factor TBX5 (T-box protein 5).			
GN	TXB5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97141914; PubMed=8988164;			
RA	Li Q.Y., Newbury-Ecob R., Tertrett J.A., Wilson D.I., Curtis A.,			
RA	Yi C.H., Bullen P.J., Strachan T., Robson S., Bonnet D.,			
RA	Young I.E., Raeburn J.A., Buckler A.J., Gebuhr T., Law D.J.,			
RA	Brook J.D.;			
RT	"Holt-Oram syndrome is caused by mutations in TBX5, a member of the			
RT	Brachyury (T) gene family.";			
RL	Nat. Genet. 15:21-29(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANT HOS GLN-237.			
RX	MEDLINE=97141915; PubMed=8988165;			
RA	Basson C.T., Bachinsky D.R., Lin R.C., Levi T., Elkins J.A.,			
RA	Souls J., Grayzel D., Kroumpouzou E., Traill T.A.,			
RA	Leplanc-Stracski J., Renault B., Kucherlapati R., Seidman J.G.,			
RA	Seidman C.E.;			
RT	"Mutations in human TBX5 cause limb and cardiac malformation in			
RT	Holt-Oram syndrome.";			
RL	Nat. Genet. 15:30-35(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Cross S.J., Ching Y.H., Armstrong-Buisseret L., Spranger S.,			
RA	Munnich A., Bonnet D., Pentinnen M., Jonveaux P., Newbury-Ecob R.,			
RA	Brook D.;			
RT	"The mutation spectrum of Holt-Oram syndrome.";			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	VARIANTS HOS ARG-80; GLN-237 AND TRP-237.			
RX	MEDLINE=99179990; PubMed=10077612;			
RA	Basson C.T., Huang T., Lin R.C., Bachinsky D.R., Werenowicz S.,			
RA	Vaglio A., Bruzone R., Quadrelli R., Lerone M., Romeo G., Silengo M.,			
RA	Pereira A., Krieger J., Mesquita S.F., Kamisago M., Morton C.C.,			
RA	Pierpont M.E.M., Muller C.W., Seidman J.G., Seidman C.E.;			
RT	"Different TBX5 interactions in heart and limb defined by Holt-Oram			
RT	syndrome mutations.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:2919-2924(1999).			
CC	FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES			
CC	REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN			
CC	LIMB PATTERN FORMATION.			
CC	SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A			
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.			

34	532.5	28.2	692	1	EOMD_XENLA	P79944 xenopus lae
35	531.5	28.1	448	1	BRA1_BRAFL	Q17134 brachiocto
36	531	28.1	433	1	BRAC_CHICK	P79777 gallus gall
37	528.5	28.0	535	1	TX21_HUMAN	Q9u117 homo sapien
38	527.5	27.9	440	1	BRA2_BRAFL	P80492 brachiocto
39	513	27.2	448	1	TX19_HUMAN	O60806 homo sapien
40	510	27.0	434	1	BRAC_HEMPU	Q25113 hemictrot
41	507	26.8	686	1	EOMD_HUMAN	O95936 homo sapien
42	505	26.7	688	1	EOMD_MOUSE	O54839 mus musculus
43	497	26.3	397	1	TBXT_CHICK	P79778 gallus gall
44	496.5	26.3	471	1	BRAC_HALRO	P56158 halocynthia
45	489	25.9	697	1	BYN_DROME	P55965 drosophila

ALIGNMENTS

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF140427; AAC00112.1; -
 EMBL: U57330; AAC53109.1; -
 HSSP: P24781; 1XBR.
 TRANSFAC: T04349; -
 MGD: MGI:102541; Tbx5.
 InterPro: IPR001699; TF_T-box.
 Pfam: PF00907; T-box; 1.
 PRINTS: PR00937; TBOX.
 SMART: SM00425; TBOX; 1.
 PROSITE: PS01283; TBOX_1; 1.
 PROSITE: PS01264; TBOX_2; 1.
 PROSITE: PS02522; TBOX_3; 1.
 Transcription regulation; DNA-binding; Nuclear protein;
 Developmental protein.
 DNA_BIND 63 238 T-BOX.
 CONFLICT 68 71 HEVG -> RSGV (IN REF. 2).
 CONFLICT 81 81 R -> G (IN REF. 2).
 SEQUENCE 518 AA: 57832 MW: 4990A802EB56624B CRC64;

Query Match 90.6%; Score 1712; DB 1; Length 518;
 Best Local Similarity 94.7%; Pred. No. 3.9e-134;
 Matches 320; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

Qy 1 MADADGFLAHTPLEDPAKDLPCDSKPESALGAPSPSPQAAFTQOGMEGKVFLE 60
 Db 1 MADTDEGFLAHTPLEDPAKDLPCDSKPESALGAPSPSPQAAFTQOGMEGKVFLE 60
 Qy 61 RELMLKFEVGTETMIITKAGRMFPYSYKVTGLNPKTKYILLMDIVPADDRHYKFDANK 120
 Db 61 RELMLKFEVGTETMIITKAGRMFPYSYKVTGLNPKTKYILLMDIVPADDRHYKFDANK 120
 Qy 121 WSVTGKAEPAKPGRLYVHPDSPATGAHMRQLVSFQKLKLTNNHLDPPFGHILNSMHKYQ 180
 Db 121 WSVTGKAEPAKPGRLYVHPDSPATGAHMRQLVSFQKLKLTNNHLDPPFGHILNSMHKYQ 180
 Qy 181 PRLHIVKADENNGFGSKNTAFCTHVPETAFIATVSYONHKITOLKIENPPFAKGRGSD 240
 Db 181 PRLHIVKADENNGFGSKNTAFCTHVPETAFIATVSYONHKITOLKIENPPFAKGRGSD 240
 Qy 241 DMELHRMSQSKYEPVYVPRSTVRQKVASNHSPPSSRALSTSSNLGSOYCENGYSGP 300
 Db 241 DLELRMSQSKYEPVYVPRSTVRHKTYSNHSPPSSRALSTSSNLGSOYCENGYSGP 300
 Qy 301 SQDLLPPNPYPLOEHSQIYHCTKRKE-----CDHPW 334
 Db 301 SQDLLPPNPYPLOEHSQIYHCTKRKEDECSSTEPY 338

RESULT 3
 TBX4_HUMAN
 ID TBX4_HUMAN STANDARD; PRT; 545 AA.
 AC P57082;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX4 (T-box protein 4).
 GN TBX4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20399572; PubMed=10945475;
 RA Yi C.H., Russ A., Brook J.D.;
 RT "Virtual cloning and physical mapping of a human T-box gene, TBX4";
 RL Genomics 67:92-95(2000).

CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
 CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
 CC LIMB PATTERN FORMATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF188703; AAF68854.1; -
 CC HSSP: P24781; 1XBR.
 TRANSFAC: T04349; -
 Genew: HGNC:11603; Tbx4.
 MIM: 601719; -
 InterPro: IPR001699; TF_T-box.
 Pfam: PF00907; T-box; 1.
 PRINTS: PR00937; TBOX.
 SMART: SM00425; TBOX; 1.
 PROSITE: PS01283; TBOX_1; 1.
 PROSITE: PS01264; TBOX_2; 1.
 PROSITE: PS02522; TBOX_3; 1.
 Transcription regulation; DNA-binding; Nuclear protein;
 Developmental protein.
 DNA_BIND 71 251 T-BOX.
 SEQUENCE 545 AA: 60216 MW: 89CB5B9F873D3FC7 CRC64;

Query Match 59.8%; Score 1130.5; DB 1; Length 545;
 Best Local Similarity 64.3%; Pred. No. 5.3e-86;
 Matches 229; Conservative 31; Mismatches 75; Indels 21; Gaps 7;

Qy 1 MADADGFLAHTPLEDPAKDLPCDSKPESALGAPSPSPQAAFTQOGMEGKVFLE 60
 Db 7 LSESEAF---RAP-GPALGEASANAPEALAPGLSGAALGSPGPGADVWAAAAAEQ 62
 Qy 50 GMEGKVFLEHERELWLKFEVGTETMIITKAGRMFPYSYKVTGLNPKTKYILLMDIVPA 109
 Db 63 TIENIKVGLHEKELWLKFEVGTETMIITKAGRMFPYSYKVTGLNPKTKYILLMDIVPA 122
 Qy 110 DDHYKFTADNKSQVTSKAEPAKPGRLYVHPDSPATGAHMRQLVSFQKLKLTNNHLDPPFG 169
 Db 123 DDHYKFTCDNKMVAGKAEPAKPGRLYVHPDSPATGAHMRQLVSFQKLKLTNNHLDPPFG 182
 Qy 170 HIILNSMHKYQPRLHIVKADENNGFGSKNTAFCTHVPETAFIATVSYONHKITOLKIEN 229
 Db 183 HIILNSMHKYQPRLHIVKADENNGFGSKNTAFCTHVPETAFIATVSYONHKITOLKIEN 242
 Qy 230 NPFAKGRGSDDMELHRMSQSKYEPVYVPRSTVRQKVASNHSPPSSRALSTSSNLGS 289
 Db 243 NPFAKGRGSDSDL-RVARLOSKEYPVIKSIWRQLISPOLSATPDVGLLTHQALQ 301
 Qy 290 QYOCENGVS---GPSQDLLPPNPYPLOEHSQIYHCTKRKEGCDHPWISFLSYL 342
 Db 302 HYOHENGAHSLAEXQDL--PLSTFTPTQRDSSLFYHCLKRDGTRHLDLPCKRSYL 355

RESULT 4
 TBX4_MOUSE
 ID TBX4_MOUSE STANDARD; PRT; 173 AA.
 AC P70325;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX4 (T-box protein 4) (Fragment).
 GN TBX4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;

QY 17 PDAKDLPCDSKPSALGAPSKSPSSQAFAFTQGMGKIVFLHRELWLKHEVGTMI 76
 Db 5 PDKF-APCD-ALASPLEP-YPOSSIT-ATLMDGLMKHFGTCTMI 49
 QY 77 TKAGRMFSPYKVTGLNPKTKYILLMDIVPADDHRYKFAADNKWSVTGKAEPAMPGR 136
 Db 50 TKSGRMFPQCKIKVSGILIPYAKVLMVLVDFVPVDFNRYKWKNDQEVAGKAEPOLPCRT 109
 QY 137 VHPDSPATCAHNMROLVSFQKLKLTNNHLDPFCHILNLSMHRYOPRLHVKADENNGFGS 196
 Db 110 VHPDSPAPGSHMKPEVSPQKLKLTNTLQHGHIILSHMRYKRFHIVAADL--FSV 167
 QY 197 KNTAFCTHVEPPTAPIAVTSYONHKITOLKIENNPFAKCFR-----GSDMELHRM 247
 Db 168 RWSIFQVSEFETVTSVATQNEQITRLADNNPFAKCFRHEGKNTTREGRAKSQKSP 227
 QY 248 SRWQKEYVPVPRSTVRK-----VASHNSFP-----SSSRALSTSSN 286
 Db 228 AKGKRLPEEKESGAERDEKENVVKEENPIVSSGYPFWISEONSSHAPPAASP 287
 QY 287 LGSQYOCENGVSQDLLLLPPNPYPPLPQEHs-----QIYHCTKRKGECDHP 333
 Db 288 APABQR-----EGPAREQOVTPSYOTYRFEAGDSQQLPSRDVAALNDFGRG-HP 338

RESULT 8

TXB2_MOUSE
 ID TXB2_MOUSE STANDARD: PRT: 701 AA.
 AC Q60707;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX2 (T-box protein 2).
 GN TBX2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL EXPRESSION.
 RC TISSUE=Embryo;
 RX MEDLINE=95004605; PubMed=7920656;
 RA Bollag R.J., Siegfried Z., Cebra-Thomas J.A., Garvey N., Davison E.M.,
 RA Silver L.M.,
 RT "An ancient family of embryonically expressed mouse genes sharing a
 RT conserved protein motif with the T locus."
 RL Nat. Genet. 7:383-389(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=10770922;
 RA Carreira S., Liu B., Goding C.R.,
 RA Chapman D.L., Garvey N., Hancock S., Alexiou M., Agulnik S.I.,
 RA Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M.,
 RA Papaioannou V.E.,
 RT "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
 RT development."
 RL Dev. Dyn. 206:379-390(1996).
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
 CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
 CC LIMB PATTERN FORMATION
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: IN ADULTS, HIGHEST LEVELS IN LUNG. ALSO FOUND
 CC IN HEART, KIDNEY, AND OVARY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION FIRST OBSERVED AT DAY 9.5 IN
 CC THE OTIC AND OPTIC VESICLES AND IN THE FACIAL REGION. AT DAY
 CC 12.5, EXPRESSED IN THE TRIGEMINAL GANGLIA, FACIAL REGIONS, RETINA

CC AND LIMB BUD MESENCHYME. IN LATER STAGES, FOUND IN EAR PINNAE,
 CC THE MILK LINE, LUNG MESENCHYME, BODY WALL, GENITAL RIDGE AND
 CC DEVELOPING NERVOUS SYSTEM.
 CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U15566; AAC52697.1;
 CC EMBL; AE244917; AAF90050.1;
 CC DR HSP; P24781; IGBR.
 CC TRANSFAC; T04358;
 CC MGD; MGI:98494; Tbx2.
 CC InterPro: IPR001699; TF_T-box.
 CC Pfam; PF00907; T-box; 1.
 CC PRINTS; PR00937; TBOX.
 CC SMART; SM00425; TBOX; 1.
 CC PROSITE; PS01283; TBOX_1; 1.
 CC PROSITE; PS01264; TBOX_2; 1.
 CC PROSITE; PS0252; TBOX_3; 1.
 CC KW Transcription regulation; DNA-binding; Nuclear protein;
 CC Developmental protein.
 CC FT DOMAIN 48 63 POLY-ALA.
 CC FT DNA_BIND 104 277 T-BOX.
 CC FT DOMAIN 572 580 POLY-ALA.
 CC FT DOMAIN 586 594 POLY-ALA.
 CC SQ SEQUENCE 701 AA; 74244 MW; 8D90ED6DA32B3859 CRC64;
 CC
 CC Query Match 34.4%; Score 649.5; DB 1; Length 701;
 CC Best Local Similarity 44.3%; Pred. No. 4.6e-46;
 CC Matches 145; Conservative 41; Mismatches 82; Indels 59; Gaps 8;
 CC
 CC Qy 23 PCDSKPESALGAPSKSP-----SSQAFAFTQGM----- 51
 CC Db 29 PALALPGALGKPLDPGLAGAAAAAAGLHVSAHGPHPAHLRLSKLSLEP 88
 CC Qy 52 ----EGIKVFLHRELWLKHEVGTMIITKAGRMFSPYKVTGLNPKTKYILLMDIV 107
 CC Db 89 DEVEDDPKVTLEAKELMDQFKLGTETVITKSGRMFPFVKRVYSGDLKAKYILLMDIV 148
 CC Qy 108 PADDHRYKFAADNKWSVTGKAEPAMPGRLYVHPDSPATGAHNMROLVSFQKLKLTNNHLD 167
 CC Db 149 AADDCRYKFNRSRMVAGKADPEMPKRMVTHPDSPATGEOMWMAKPVAFHKLKLTNNISDK 208
 CC Qy 168 FGHILNLSMHRYOPRLHVKADENNGFGSKNTAFCTHVFPEAFIAYTSYONHKITOLKI 227
 CC Db 209 HGFTILNLSMHRYOPRHHVRA--NDILKLPYSTERTYVFPETDFIAYTAYQNDKITOLKI 266
 CC Qy 228 ENNPFAKFRGSDMELHRMSRMQSKKEYPVVPRSTVRKQVSNHISPFSSSRALSTSSNL 287
 CC Db 267 DNNPFAKGFRTGN-----GRREKRKQLTLP--TLR--LYEEHCKPGRDCAESDASS-- 314
 CC Qy 288 GSOYOCENGVSQDLLLLPPNPYPPLP 314
 CC Db 315 ----CD-----PPAREPPSPSAAP 331
 CC
 CC RESULT 9
 CC TXB6_HUMAN
 CC ID TXB6_HUMAN STANDARD: PRT: 436 AA.
 CC AC O5947;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE T-box transcription factor TBX6 (T-box protein 6).
 CC GN TBX6.
 CC OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FT DOMAIN 48 63 POLY-ALA.
 FT DNA_BIND 104 277 T-BOX-GLY.
 FT DOMAIN 507 517 POLY-GLY.
 FT DOMAIN 571 579 POLY-ALA.
 FT DOMAIN 585 593 POLY-ALA.
 FT CONFLICT 155 168 Y
 FT CONFLICT 165 168 AGKA - (IN REF. 2).
 FT CONFLICT 165 168 AGKA - (IN REF. 2).
 SQ SEQUENCE 702 AA: 74194 MW: C6477134C69D7C2C CRC64;
 Query Match 33.5%; Score 633.5; DB 1; Length 702;
 Best Local Similarity 52.0%; Pred. No. 9.6e-45;
 Matches 130; Conservative 30; Mismatches 53; Indels 37; Gaps 3;
 QY 23 PCDSKPESALGAPSKSP-----SSPOAAFTQCGM----- 51
 DB 29 PALALPPGALAKLPDPDLGACAAAAAAGLHVSLGPHPPAAHLRSLSLEPE 88
 QY 52 ----EGIKVFLHERELWLFKEVGTETIITKAGRRMFPSPYKVTGLNPKTKYILLMDIV 107
 DB 89 DEVEDDPKVTLEAKELMDQPHKLGTEMTVTKSGRRMFPFPKVRVSGLDKAKYILLMDIV 148
 QY 108 PADDHYKFEADNKSVCCKAPMPGRGLYVHPDSPATGAHNRQLVSPFKLKTNNHLD 167
 DB 149 AADCRKFKHNSRMVAVGADKDPENPKMYIHPDSPATGQWMAKPVAFHKLKTNNISDK 208
 QY 168 FGHIIILSMHKYQPRHLIVKADENNGFGSKNTAFCTHVEPETAFTAVTSYONHKITOLKI 227
 DB 209 HGFTILSMHKYQPRHLIVRA--NDILKLPYSTERYVEPEPETAFTAVTAYONDKITOLKI 266
 QY 228 ENNPFAKGR 237
 DB 267 DNNPFAKGR 276

RESULT 11

ID TBX6_MOUSE STANDARD; PRT; 540 AA.
 AC P70327;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX6 (T-box protein 6).
 GN TBX6
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutaria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=97032942; PubMed=9878690;
 RA Aguilnik S.I., Garvey N., Hancock S., Ruvinsky I., Chapman D.L.,
 RA Aguilnik I., Bolleg R.J., Papaioannou V.E., Silver L.M.;
 RA "Evolution of mouse T-box genes by tandem duplication and cluster
 dispersion";
 FT Genet. 144:249-254 (1996).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Gastrula;
 RX MEDLINE=97115702; PubMed=9954725;
 RA Chapman D.L., Aguilnik I., Hancock S., Silver L.M., Papaioannou V.E.;
 RA "Tbx6, a mouse T-box gene implicated in paraxial mesoderm formation at
 gastrulation";
 FT Dev. Biol. 180:534-542 (1996).
 RL [3]
 RP FUNCTION
 RX MEDLINE=98140705; PubMed=9490412;
 RA Chapman D.L., Papaioannou V.E.;
 RA "Three neural tubes in mouse embryos with mutations in the T-box gene
 Tbx6";
 FT Nature 391:695-697 (1998).
 RL Nature 391:695-697 (1998).
 CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
 DEVELOPMENTAL PROCESSES. REQUIRED FOR SPECIFICATION OF PARAXIAL

CC MESODERM STRUCTURES DURING GASTRULATION. IN ITS ABSENCE CELLS
 CC DESTINED TO FORM POSTERIOR SOMITES DIFFERENTIATE ALONG A NEURONAL
 CC PATHWAY.
 CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -!- DEVELOPMENTAL STAGE: TBX6 IS FIRST DETECTED IN THE GASTRULATION
 CC STAGE IN THE PRIMITIVE STREAK AND NEWLY RECRUITED TO PARAXIAL
 CC MESODERM. LATER IN DEVELOPMENT IT IS RESTRICTED TO PRESOMITIC,
 CC PARAXIAL MESODERM AND TO THE TAIL BUD, WHICH REPLACES THE STREAK
 CC AS THE SOURCE OF MESODERM.
 CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: U57331; AAC53110.1;
 CC HSSP: P24781; 1XBR;
 CC TRANSFAC: T04417;
 CC MGI: MGI:102639; Tbx6
 CC InterPro: IPR001699; TF_T-box.
 CC Pfam: PF00507; T-box; 1.
 CC PRINTS: PR00937; TBOX.
 CC SMART: SM00435; TBOX; 1.
 CC PROSITE: PS01283; TBOX_1; 1.
 CC PROSITE: PS01264; TBOX_2; 1.
 CC PROSITE: PS0252; TBOX_3; 1.
 CC TRANSFAC: PS0252; TBOX_3; 1.
 CC Transcription regulation; DNA-binding; Nuclear protein;
 CC Developmental protein;
 CC DOMAIN 6 84 POLY-ALA.
 CC DOMAIN 70 84 POLY-PRO.
 CC DNABIND 100 273 T-BOX.
 CC SEQUENCE 540 AA: 58628 MW: BC834CE2745E8E61 CRC64;
 Query Match 33.3%; Score 629.5; DB 1; Length 540;
 Best Local Similarity 43.4%; Pred. No. 1.5e-44;
 Matches 138; Conservative 45; Mismatches 80; Indels 55; Gaps 9;
 QY 11 AHPTLEPDADKDLPCDSKPSALGAPSKSPSSPOAAFTQCGMEGIKVFLHERELWLFKEV 70
 DB 62 AAALPFS-----PIALG-PETAPPPPEAL---HSLPGVSLNQLWKEFSAV 107
 QY 71 GTEMIITKAGRRMFPSPYKVTGLNPKTKYILLMDIVPADDHYKFEADNKSVCCKAP 130
 DB 108 GTEMIITKAGRRMFPSPYKVTGLNPKTKYILLMDIVPADDHYKFEADNKSVCCKAP 167
 QY 131 MFGRLYVHPDSPATGAHNRQLVSPFKLKTNNHLDPFHIIILSMHKYQPRHLIVKADE 190
 DB 168 LPDRYIHPDSPATGAHNRQLVSPFKLKTNNHLDPFHIIILSMHKYQPRHLIVRA-- 225
 QY 191 NNGFGSKNTAFCTH-----VFPETAFTAVTSYONHKITOLKIENNPFAKGR-GSDD 241
 DB 226 -----TOLCSOHNGGVASFPEPTFTISVTAYONPRITOLKIENNPFAKGRNGRN 277
 QY 242 MELHMSRMQSKPEYVVPVPRSTVTRQKVASNHSFPSESRLSTSNLSGQYQCENGVSPGS 301
 DB 278 CKREDAKVRKRLGPEPVAT--EACGSGDTPGG-----PCDSTLGGDI 319
 QY 302 QDLPPPPNPPY-----LPQ 315
 DB 320 RDSDEQAPTPGKLLLPQ 337
 RESULT 12
 ID TBX1_HUMAN STANDARD; PRT; 398 AA.
 AC O43435; O43436;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT *The genome sequence of *Drosophila melanogaster*.;
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
 RC TISSUE-Larva;
 RX MEDLINE=93261414; PubMed=8492800;
 RA Poock B., Balles J., Pflugfelder G.O.;
 RT *Transcript identification in the optomotor-blind locus of *Drosophila*
 RT melanogaster by intragenic recombination mapping and PCR-aided
 RT sequence analysis of lethal point mutations.;;
 RL Mol. Gen. Genet. 238:325-332(1993).
 CC -1- FUNCTION: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
 CC REGULATOR. FLIES WITH L(1)LOMB MUTATIONS SHOW SEVERE MALDEVELOPMENT
 CC OF THE OPTIC LOBES, REDUCTION IN WING SIZE AND AN INCREASED
 CC ABDOMINAL PIGMENTATION. THEY DIE DURING THE PUPAL STAGE.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: IN THIRD-INSTAR LARVAE IT IS FOUND IN THE
 CC BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
 CC THE THORACIC PART OF THE VENTRAL GANGLION.
 CC -1- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
 CC EMBRYOGENESIS, THE SECOND DAY OF PUPAL DEVELOPMENT AND IN THE
 CC ADULT.
 CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC *This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M81796; AAA28736.1; .
 DR EMBL; AE003431; AAF45946.1; .
 DR EMBL; S61732; AAB36697.1; JOINED.
 DR EMBL; S61727; AAB36697.1; JOINED.
 DR EMBL; S61749; AAB36697.1; JOINED.
 DR EMBL; S61744; AAB36699.1; JOINED.
 DR EMBL; S61955; AAB36699.1; JOINED.
 DR HSP; P24781; IYBR.
 DR F1Base; FB46000179; b1.
 DR InterPro; IPR001639; TF-T-box.
 DR Pfam; PF00907; T-box.
 DR PRINTS; SM00327; TBOX.
 DR SMART; SM00327; TBOX.
 DR PROSITE; PS01283; TBOX_1; 1.
 DR PROSITE; PS01283; TBOX_2; 1.
 DR PROSITE; PS0282; TBOX_3; 1.
 KW DNA-binding; Nucleic acid protein; Transcription regulation.
 FT DOMAIN 51; 107.
 FT DOMAIN 101; 107.
 FT DOMAIN 179; 184.
 FT DOMAIN 229; 236.
 FT DOMAIN 238; 244.
 FT DNA-BIND 332; 513.
 FT DOMAIN 574; 577.
 FT DOMAIN 607; 692.
 FT DOMAIN 823; 916.
 FT DOMAIN 926; 956.
 FT CONFLICT 10; 90.
 FT CONFLICT 216; 216.
 FT CONFLICT 511; 511.
 FT CONFLICT 823; 823.
 FT CONFLICT 976; 988.
 SQ SEQUENCE 988 AA; 103992 MW; 032B7A471743FC9 CRC64;
 Query Match 33.1%; Score 626; DB 1; Length 988;
 Best Local Similarity 47.1%; Pred. No. 6.4e-44;

Matches 136; Conservative 38; Mismatches 69; Indels 46; Gaps 6;
 QY 55 KYFLHERLWLFKEVGTETMIITKAGRMFPYSKYVKTGLNPKTKYILLMDIVPADDDHY 114
 DB 329 KVTLEKDLWEKFKHGLTETWITKSGRPFQPMKFRVSGLDKAKYKAKYILLDIDVAADRY 388
 QY 115 KFAADKNSVTGKABPMPGRGLYVHPDSPAHAHWRQLVSKFQKLLTNHLDLPFGHI--- 171
 DB 389 KFHSKRWVAGKADPEKPKMYTHPDSPPTGQMMQKVVVSHKLLKLNNTSKIGFVSTT 448
 QY 172 ILNSHKYQPRHLVHKADENNGSGKNTAFCTHVPETAFATVSYQNHKITOLKIENNP 231
 DB 449 ILNSHKYQPRHLVRA--NDILKLPYSTFTYVKEFEFTAVTAYQNEKITOLKIDNNP 506
 QY 232 FAKGFR--GSDMELHMRMSQSKYVPRSTVTKVASHNSPSSSESSALSTSSNLGS 289
 DB 507 FAKGFRDTCAGKRE-----KNCYQALMSNR--GSDOKLNPHTVSSS 547
 QY 290 QYQENGSGVSGQDLLP-----PPNPYPLPOEHS 318
 DB 548 RAPLHLGHAGRPPLHPLHAAALLDQDDDDKLLDVGVGPPSPPLPLSHS 596
 RESULT 14
 ID TEX3_HUMAN STANDARD: PRT; 742 AA.
 AC O15119; O9UKF8;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX3 (T-box protein 3).
 GN TBX3.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM I), AND FUNCTION.
 RC TISSUE-Breast carcinoma;
 RX MEDLINE=9598688; PubMed=10468588;
 RA "de M.-L., Wen L., Campbell C.E., Wu J.Y., Rao Y.;
 RT Transcription repression by *Xenopus* Et and its human ortholog TBX3, a
 RT gene involved in ulnar-mammary syndrome.
 RL Proc. Natl. Acad. Sci. U.S.A. 96:10212-10217(1999).
 RP [2]
 RP SEQUENCE OF 1-488 FROM N.A. (ISOFORM I).
 RC TISSUE-Kidney;
 RX MEDLINE=97351519; PubMed=9207801;
 RA Bamshead M., Lin R.C., Law D.J., Watkins W.S., Krakowiak P.A.,
 RA Moore M.E., Franceschini P., Lala R., Holmes L.B., Gebuhr T.C.,
 RA Schinzel A., Bruneau B.G., Seidman J.G., Seidman C.E., Jorde L.B.;
 RT Mutations in human TBX3 alter limb, apocrine and genital development
 RT in ulnar-mammary syndrome;
 RL Nat. Genet. 16:311-315(1997).
 RP [3]
 RP SEQUENCE OF 591-742 FROM N.A. ALTERNATIVE SPLICING, AND VARIANTS UMS.
 RX MEDLINE=99264236; PubMed=10330342;
 RA Bamshead M., Le T., Watkins W.S., Dixon M.E., Kramer B.E., Roeder A.D.,
 RA Carey J.C., Root S., Schinzel A., Van Maldergem L., Gardner R.J.M.,
 RA Lin R.C., Seidman C.E., Seidman J.G., Wallerstein R., Moran E.,
 RA Stephen R., Campbell C.E., Jorde L.B.;
 RT The spectrum of mutations in TBX3: genotype/phenotype relationship in
 RT ulnar-mammary syndrome.;;
 RL Am. J. Hum. Genet. 64:1550-1562(1999).
 RP [4]
 RP SEQUENCE FROM N.A. (ISOFORM III).
 RC TISSUE-Adrenal gland;
 RX Song H., Gao G., Feng Y., Ren S., Chen Z., Han Z.;
 RT A novel gene expressed in human adrenal gland
 RT "A novel gene expressed in human adrenal gland"
 RL Submitted (Dec-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN DEVELOPMENTAL
 CC PROCESSES. PROBABLY PLAYS A ROLE IN LIMB PATTERN FORMATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC

Query Match 32.6%; Score 616.5; DB 1; Length 251;
Best Local Similarity 53.5%; Pred. No. 6.2e-44;
Matches 122; Conservative 34; Mismatches 55; Indels 17; Gaps 4;

QY	33	CAPSKSPSPQAA-----FTQGMGKIKVLFHERELWLKPFHEVGTEMIITKAGRR	82
DB	24	CGSGSPSPSSSLCTEPLIPTPIIPEEMAKIACSLKTELWDKPFHELCTEMIITKSGRR	83
QY	83	MPPSYKVKVTGLNPKTKYILLMDIVPADDDHRYKPADNK--WSVTGKAEPAMFGRLYVHPD	140
DB	84	MPPTIRVSFSGVDPEAKYIVLMDIVPDVNDKRYRYAYHRSSWLAVAGKADPPLPARLYVHPD	143
QY	141	SPATCAHMRQLVSFQRLKLTNNHLDPFCHILNSMHKYQPRHLHIVKADENNG--FGSKN	198
DB	144	SPFTGEQLLKQVSEFKVKLTNNELDQHGHIILNSMHKYQPRVHIKKKDHDTASLLNLKS	203
QY	199	TAFCTHVFPETAFIYTSYONHKITQLKIENNPFAKFGSGD---DME	243
DB	204	EEFRFTFPETVFTAVTAYONQLITKLKIDSNPFAGFRDSSRLTDIE	251

Search completed: June 17, 2003, 09:47:36
Job time : 13 secs


```
QY 61 RELMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPADHRYKFEADNK 120
DB 78 RELMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPADHRYKFEADNK 137
QY 121 WSVTGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCHIIILNSMHRKYQ 180
DB 138 WSVTGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCHIIILNSMHRKYQ 197
QY 181 PRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENNPFKFRGSD 240
DB 198 PRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENNPFKFRGSD 257
QY 241 DMEIHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSOYOCENGVS GP 300
DB 258 DMEIHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSOYOCENGVS ST 317
QY 301 SODLLPPNPYPLPOEHSOIYHCTKRKE---CDHPW 334
DB 318 SODLLPPTNPYPLPOEHSOIYHCTKRKEDECSSTTEHPY 355
```

RESULT 2

```
Q9PWE8 PRELIMINARY: PRT: 521 AA.
AC 09PWE8:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE T-box transcription factor Tbx5.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC NCBI_TaxID=9031.
RN [1]
RP MEDLINE FROM N.A.
RX MEDLINE-99303540; PubMed-10373308;
RA Bruner B.G., Logan M., Davis N., Levi T., Tablin C.J., Seidman J.G.,
RA Seidman C.E.;
RT "Chamber-specific cardiac expression of Tbx5 and heart defects in
RT Holt-Oram syndrome."
RL dev. Biol. 211:100-108(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Logan M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR HSPF: P24781.1 JXBR.
DR TRANSPAC: T04348.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PF00907; T-box.
DR PRINTS: PR00937; TBOX.
DR SMART: SM00425; TBOX.1.
DR PROSITE: PS01283; TBOX_1.1.
DR PROSITE: PS01264; TBOX_2.1.
DR PROSITE: PS50252; TBOX_3.1.
SQ SEQUENCE 521 AA; 58402 MW; 17633D0F1DCE30B CRC64;
```

Query Match 85.3%; Score 1611; DB 13; Length 521;
Best Local Similarity 88.8%; Pred. No. 3.8e-138;
Matches 300; Conservative 15; Mismatches 19; Indels 4; Gaps 1;

```
QY 1 MADADEGFIATPRPPADKDLPCDSKPESALGAPSPSPSPDPAFTOOGMGIKVFLE 60
DB 1 MADTEBEGFGLSPPTVDSSEKELQAAKODPDLGLTTSKAPTSPQAFTQCGMEGIKVLE 60
QY 61 RELMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPADHRYKFEADNK 120
DB 61 RELMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPADHRYKFEADNK 120
QY 121 WSVTGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCHIIILNSMHRKYQ 180
```

```
DB 121 WSVTGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCHIIILNSMHRKYQ 180
QY 181 PRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENNPFKFRGSD 240
DB 181 PRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENNPFKFRGSD 240
QY 241 DMEIHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSOYOCENGVS GP 300
DB 241 DMEIHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSOYOCENGVS ST 300
QY 301 SODLLPPNPYPLPOEHSOIYHCTKRKE---CDHPW 334
DB 301 SODLLPPTNPYPLPOEHSOIYHCTKRKEDECSSTTEHPY 338
```

RESULT 3

```
Q96TB0 PRELIMINARY: PRT: 342 AA.
AC 096TB0:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Transcription factor T-box 5.
GN TBX5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Demura M., Yoneda T., Takeda Y., Furukawa K., Mabuchi H.;
RT "Human transcription factor TBX5 mRNA, alternatively spliced."
RL EMBL: AB051068; BAB55448.1.
DR EMBL: AB051068; BAB55448.1.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PF00907; T-box.
DR PRINTS: PR00937; TBOX.
DR PROSITE: PS01283; TBOX_1.
DR PROSITE: PS01264; TBOX_2.
DR PROSITE: PS50252; TBOX_3.
SQ SEQUENCE 342 AA; 38916 MW; C368AD992EDPFD4A CRC64;
```

Query Match 79.6%; Score 1503; DB 4; Length 342;
Best Local Similarity 97.6%; Pred. No. 1.4e-128;
Matches 281; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

```
QY 51 MEGIKVFLHERMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPAD 110
DB 1 MEGIKVFLHERMLKHEVGTMTTKAGRMFSPYKVTGTLNPKTKIILMDIVPAD 110
QY 111 DHRKFPADNKMSVTCGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCH 170
DB 61 DHRKFPADNKMSVTCGAERAMPGRLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPFCH 120
QY 171 IILNSMHRKYQPRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENN 230
DB 121 IILNSMHRKYQPRHIVADENNFGSKNTAFCTHVPETAFIAVTSYONHKITOLKIENN 180
QY 231 PFAKFRGSDMELHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSO 290
DB 181 PFAKFRGSDMELHRMSRMOQKEYPVVPRSTVRQKVASNHSPPSSSESRALSTSSNLGSO 240
QY 291 YOCENGVSQSDLLPPNPYPLPOEHSOIYHCTKRKE---CDHPW 334
DB 241 YOCENGVSQSDLLPPNPYPLPOEHSOIYHCTKRKEDECSSTTEHPY 288
```

RESULT 4

```
Q9W7C2 PRELIMINARY: PRT: 519 AA.
AC 09W7C2:
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
```


Query Match	79.5%;	Score 1501.5;	DB 13;	Length 519;
Best Local Similarity	86.0%;	Pred. No. 3.5e-128;		
Matches	282;	Conservative	16;	Mismatches 29;
			Indels	1;
			Gaps	1

QY	1	MADDEGGGLNHTLEPAPKDLPCDSKREESLAPVSKSPSSQOAAFTQOQMGIVFYLHE	60
QY	1	: : : : : : : : : : : : : : :	60
Db	1	MAOTEEAYGMBDTVEAPRFKELOCEBPQDNQIGASSKFTPTSPRAFTQOQMGIVFLHE	60
QY	61	RELWLKFLFVETEMITITAGRRMRPSSYKYVXTGTLNPKTKYILLMDIVADDRHYEFADNK	120
Db	61	RELWLKFLFVETEMITITAGRRMRPSSYKYVXTGTLNPKTKYILLMDIVADDRHYEFADNK	120
QY	121	WSYTGKAEAPAPGRLYVHPDSPAIGAHMMROLVYSFQKLTINHLIDPRGHIILNSMHKYO	180
Db	121	WSYTGKAEAPAPGRLYVHPDSPAIGAHMMROLVYSFQKLTINHLIDPRGHIILNSMHKYO	180
QY	181	PRLLIVVADENNNGSGSKTAFCTVYFPETALAYTSYONHKITOLKIENTNPPAKCFRSD	240
Db	181	PRLLIVVADENNNGSGSKTAFCTVYFPETALAYTSYONHKITOLKIENTNPPAKCFRSD	240
QY	241	DMELHRMSRMOSKEYPVVRSTVROKVASNHSPSSFEESRALSTSSNLTQYQCEGVSGP	300
Db	241	DMELHRMSRMOSKEYPVVRSTVROKVASNHSPSSFEETRINTGSSLTNSQYQCEGVSGT	300
QY	301	SQDLLPPRPNTY-PLRQEHSLQIYHCTKKR 327	
Db	301	SQDLLPPSSAYTSLPHESGTIIYHCTKKR 328	

RESULT	5		
09PUS7			
ID	09PUS7	PRELIMINARY;	PRT: 485 AA.
AC	09PUS7		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	T-box transcription factor Tbx5.		
GN	TBX5.		
OS	Brachydanio rerio (zebrafish) (zebra danio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCHI_TaxID=7955;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=99425182; PubMed=10495283;		
RA	Tamura K., Yonel-Tamura S., Belmonte J.C.;		

Query Match	73.7%	Score 1392:	DB 13:	Length 485:
Best Local Similarity	78.5%	Pred. No. 3e-118:		
Matches 267:	Conservative 24:	Mismatches 33:	Indels 10:	Gaps 5:

```

QY 1 MADDDEFGFGLAHFPLEPDADKLDPDSDSPESALGAPSPSPSSPQAAFPQOQMEGJIKVFLHE 60
    1111111111111111111111111111111111111111111111111111111111111111
Db 1 MADDSEFFRLQNSPDSSEPKDLQNEGSKDKQMAVSKSPSS--QYTTYIQQGEKGIKVLHE 59
QY 61 RELMLKHEHVECTEMITTKAGRMFPRSKVYKUTGLNPKTKYLLMDIYRPAHDHRYKFFADNK 120
    1111111111111111111111111111111111111111111111111111111111111111
Db 60 RELMTAKHEVGETEMITTKAGRMFPRSKVYKUTGLNPKTKYLLMDIYRPAHDHRYKFFADNK 119
QY 121 WSYVGAKEAPAMPGRILYVHPDSPATGAMHMROLVASFQKILKLTNNHIDPFGHITLMSMKRYQ 180
    1111111111111111111111111111111111111111111111111111111111111111
Db 120 WSYVGAKEAPAMPGRILYVHPDSPATGAMHMROLVASFQKILKLTNNHIDPFGHITLMSMKRYQ 179
QY 181 PRILIIYVADENNGSGSKNTACTCTVFPETATIAVTSQONKRTIOLKTIENNPFAAGFPGSD 240
    1111111111111111111111111111111111111111111111111111111111111111
Db 180 PRILIIYVADENNGSGSKNTACTCTVFPETATIAVTSQONKRTIOLKTIENNPFAAGFPGSD 239
QY 241 DMEIHRMSRMOS--KEYVPVPRSTYRQVAVNSHPSESSRALTSTSNLGSYOQKGVSSG 299
    1111111111111111111111111111111111111111111111111111111111111111
Db 240 DMEIHRMSRMOSTKEYVPVPRSTYRQVAVNSHPSESSRALTSTSNLGSYOQKGVSSG 299
QY 300 PSODILPPNPYPLPOEHSIOLYHCTKKR--GEC---DHPM 334
    1111111111111111111111111111111111111111111111111111111111111111
Db 300 TSODILPPSSSY---HEHTDYHCTIKRVEDECPAGEHPY 336

```

RESULT 6		
Q91AK8		
ID	Q91AK8	PRELIMINARY; PRT; 492 AA.
AC	Q91AK8;	
DT	01-OCT-2000 (TREMBLrel. 15, created)	
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)	
DT	01-MAR-2002 (TREMBLrel. 20, last annotation update)	
DE	T-box transcription factor tbx5.	
OS	Brachydanio rerio (zebrafish) (Zebra danio).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7953;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=20130955; PubMed=1066415;	
RA	Rutinsky I., Oates A.C., Silver E.M., Ho R.K.;	
RT	"The evolution of paired appendages in vertebrates: T-box genes in the zebrafish.";	
RL	Dev. Genes Evol. 210:82-91(2000).	
DR	EMBL: AF179407; AAF59837.1; -.	
DR	HSSP: P24781; 1XR	
DR	InterPro: IPR001899; Gram_pos_anchor.	
DR	InterPro: IPR001699; TF_T-box.	
DR	Pfam: pf00907; T-box; 1.	
DR	PRINTS: PR00937; TBOX.	
DR	SMART: SM00425; TBOX; 1.	
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.	
DR	PROSITE: PS01283; TBOX_1; 1.	
DR	PROSITE: PS01264; TBOX_2; 1.	

```

DR PROSITE: PSS0252; TBOX_3; 1.
SQ SEQUENCE 492 AA; 55236 MW; 5C6940D1942D2750 CRC64;
Query Match 72.4%; Score 1387; DB 13; Length 492;
Best Local Similarity 78.2%; Pred. No. 8.6e-118;
Matches 206; Conservative 24; Mismatches 40; Indels 10; Gaps 5;
OY 1 MADADGCGLAHTPELPDAKDLPCDSKPESALCAPSKSPSSPOAFTQOGMEGICIKVFLHE 60
DB 1 MADSDTRFLONSDSDPEKDLONEGSKDPKNAALSKSPSS-OTTYIQOGMEGICIKVFLHE 59
OY 61 RELMLKTFHEVGTETITKAGRRMPSYKVKVTGLNPKTKYLLMDIYPADHRYKFAADNK 120
DB 61 RELMLKTFHEVGTETITKAGRRMPSYKVKVTGLNPKTKYLLMDIYPADHRYKFAADNK 119
OY 121 WSVTGAEPAPAPGALYHPDSPATGAHMRQLVSFOKLKLTNNHLDPFCHIILNSMHKYO 180
DB 120 WSVTGAEPAPAPGALYHPDSPATGAHMRQLVSFOKLKLTNNHLDPFCHIILNSMHKYO 179
OY 181 PRLATYKADENNGGSKNTAFCTHYEPETAFIAYTSTONKKTQIKTENNPFKGRGSD 240
DB 180 PRLATYKADENNGGSKNTAFCTHYEPETAFIAYTSTONKKTQIKTENNPFKGRGSD 239
OY 241 DMELHRMSRMOS-KEYPVVPRSTVRKOVASNHSPSSSRALSTSSMLGSOYOCENGVS 299
DB 240 DMELHRMSRMOSTKEYPVVPRSTVRKOVASNHSPSSSRALSTSSMLGSOYOCENGVS 299
OY 300 PSODLLPPNPYPLPOEHSQIYCTKRK--GEC--DHPW 334
DB 300 PSODLLPPNPYPLPOEHSQIYCTKRK--GEC--DHPW 334
OY 300 TSODLLPOSSSY--HEHTQDYHCIRKRYEDECPAGEHPY 336
DB 300 TSODLLPOSSSY--HEHTQDYHCIRKRYEDECPAGEHPY 336

RESULT 7
OYPTK3 PRELIMINARY; PRT; 372 AA.
ID OYPTK3;
AC 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE T-box containing transcription factor Tbx5.1 (Fragment).
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7935;
RN 111
RP SEQUENCE FROM N.A.
RA MEDLINE:20108594; PubMed:10640716;
RA Begemann G., Ingram P.W.;
RT Developmental regulation of Tbx5 in zebrafish embryogenesis.;
RL Mech. Dev. 90:299-304(2000).
DR EMBL: AF165283; AAF22296.1; -
DR HSSP: P24781; 1XBR.
DR ZFIN: ZDB-GENE-991124-7; tbx5.
DR InterPro: IPR001699; TF_T-box.
DR Pfam: PF00907; T-box; 1.
DR PRINTS: PR00937; TBOX.
DR SMART: SM00423; TBOX_1; 1.
DR PROSITE: PS01283; TBOX_1; 1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
FT NON_TER 1
FT 372
SQ SEQUENCE 372 AA; 41764 MW; 6FDF0DE3C5340EC CRC64;
Query Match 71.4%; Score 1349; DB 13; Length 372;
Best Local Similarity 84.0%; Pred. No. 1.7e-114;
Matches 258; Conservative 15; Mismatches 24; Indels 10; Gaps 5;
OY 34 AFSKSPSPQAAFTQOGMEGICIKVFLHERELMLKTFHEVGTETITKAGRRMPSYKVKYTG 93
DB 4 AYSKSPSS-OTTYIQOGMEGICIKVFLHERELMLKTFHEVGTETITKAGRRMPSYKVKYTG 62

```

```

OY 94 LNPRTKYLLMDIYPADHRYKFAADNKSVTGAEPAPGRKLYVHDSPATGAHMRQLV 153
DB 63 LNPRTKYLLMDIYPADHRYKFAADNKSVTGAEPAPGRKLYVHDSPATGAHMRQLV 122
OY 154 SFORKLTNNHLDPFCHIILNSMHKYOPLHIVKADENNGGSKNTAFCTHYEPETAFIA 213
DB 123 SFORKLTNNHLDPFCHIILNSMHKYOPLHIVKADENNGGSKNTAFCTHYEPETAFIA 182
OY 214 VTSYONKKTQIKTENNPFKGRGSDDMELHRMSRMOS-KEYPVVPRSTVRKOVASNH 272
DB 183 VTSYONKKTQIKTENNPFKGRGSDDMELHRMSRMOSTKEYPVVPRSTVRKOVASSOS 242
OY 273 PFSSSRALSTSSMLGSOYOCENGVSGPSODLLPPNPYPLPOEHSQIYCTKRK--GEC 330
DB 243 PFSGDVGSLATGAALSSOYSCENGVSSTSDLLPOSSSY--HEHTQDYHCIRKRYEDE 299
OY 331 --DHPW 334
DB 300 PAGEHPY 306

RESULT 8
OY3288 PRELIMINARY; PRT; 541 AA.
ID OY3288;
AC 093288;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE T-box protein 4 (T-box transcription factor Tbx4).
GN Tbx4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
RA MEDLINE:98220375; PubMed:9550719;
RA Issac A., Rodriguez-Esteban C., Ryan A., Altshuler M., Tsukui T.,
RA Patel K., Tickle C., Izpisua-Belmonte J.-C.;
RT Tbx genes and limb identity in chick embryo development.;
RL Development 125:1867-1875(1998).
RN 121
RP SEQUENCE OF 66-249 FROM N.A.
RA Logan M., Simon H.-G., Tabin C.;
RT "Differential regulation of T-box transcription factors suggests a
RT role in controlling chick limb-type identity.";
RL Development 0:0-0(1998).
DR EMBL: AF033670; AAC41298.1; -
DR HSSP: AF069395; AAC23682.1; -
DR HSSP: P24781; 1XBR.
DR InterPro: IPR001699; TF_T-box.
DR Pfam: PF00907; T-box; 1.
DR PRINTS: PR00937; TBOX.
DR SMART: SM00425; TBOX_1; 1.
DR PROSITE: PS01283; TBOX_1; 1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
SQ SEQUENCE 541 AA; 60381 MW; 5744194A87CC22EB CRC64;
Query Match 61.0%; Score 1152.5; DB 13; Length 541;
Best Local Similarity 64.4%; Pred. No. 2.2e-96;
Matches 230; Conservative 33; Mismatches 79; Indels 15; Gaps 6;
OY 1 MADADGCGLAHTPELPDAKDLPCDSKPESALCAPSKSPSSPOAFTQOGMEGICIKVFLHE 59
DB 7 LSEDEGCFAPAPGADSS---AGSPVLGAVAGSSTPLSSPOLPDEQETENIKVYLH 62
OY 60 ERELMLKTFHEVGTETITKAGRRMPSYKVKVTGLNPKTKYLLMDIYPADHRYKFAADNK 119
DB 63 ERELMLKTFHEVGTETITKAGRRMPSYKVKVTGMNPKTKYLLMDIYPADHRYKFCDN 122

```

```

QY 120 KMSVTKAPAMPGRLYVHPDSFAGAMHMOVLVSFOKLTKLTNNHDPFGHILNLSMHKY 179
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 123 KMMVAKAEPAMPGRLYVHPDSFAGAMHMOVLVSFOKLTKLTNNHDPFGHILNLSMHKY 182
    || || || || || || || || || || || || || || || || || || || || || || ||
QY 180 QPRLHIVKADENNNGFSGSKNTAFCTHVPEFTAFIATVSYONKRIQOLKIENPFAKGRGS 239
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 183 QPRLHIVKADENNNAFSGSKNTAFCTHVPEFTSIVSYONKRIQOLKIENPFAKGRGS 242
    || || || || || || || || || || || || || || || || || || || || || || ||
QY 240 DQMLHHRMSRMSKEPPVPRKSTVQKVASNHSPSS--ESRALSTSSNLSOYQCENG- 296
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 243 DQSDLL-RVARIQSKKEPVYSIKIMQRIQVSSHGOLSTPDPVNPPLHSGHOTLQNHQYENGA 301
    || || || || || || || || || || || || || || || || || || || || || || ||
QY 297 ----VSGPADDLLPPNPYLLPQESHSOYLHCYKRGCEGDHWMSICPLSYFLSLGWS 349
    || || || || || || || || || || || || || || || || || || || || || || ||
Db 302 HMQFAASDTQDL--SLNTFTAQRDGSLYTHCLIKRRESRHLDDPCKRSYLEASSVVG 356
    || || || || || || || || || || || || || || || || || || || || || || ||

```

RESULT	9		
08R5F6			
ID	08R5F6	PRELIMINARY;	PRT: 552 AA.
AC	08R5F6;		
DT	01-JUN-2002 (Tremblrel, 21, Created)		
DT	01-JUN-2002 (Tremblrel, 21, Last sequence update)		
DT	01-JUN-2002 (Tremblrel, 21, Last annotation update)		
DE	T-box 4.		
GN	Tbx4.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxId=10090;		
RN	(1)		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=129/SVEV;		
RC	Arruda E.P., Bruneau B.G.;		
RA	"Mouse Tbx4 in limb development."		
RT	Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.		
RL	EMBL: AY075134; AAL77209.1.		
DR	EMBL: AY075134; AAL77209.1.		
SO	SEQUENCE 552 AA; 61142 MW; B86AB8A55E736100 CRC64;		

Query Match	60.58;	Score 1143;	DB 11;	Length 552;
Best Local Similarity	63.58;	Pred. No. 1.7e-95;		
Matches 233;	Conservative 31;	Mismatches 67;	Indels 36;	Gaps 8

QY	1	MADADGFCPLATPLERPAKODLPCDSKRESNLAGPASKP----	SSP-----	-QAFT	47
		7	LSSEBAP--RAQCPALGEASNTSTTNAEPPLATPGLSGALSSPQOGADVAAAAA		65
QY	48	OQMEGICKYFELHERELTYLFEHVEGTEMTITKAGRMFPSTYUKYTGANPKTYILLDIY			107
		66	DETENIKTKGHEKELMKKFNHAGLEMTITTAAGRMFPSTYUKYTGANPKTYILLDIY		125
Db	108	PADNHYKFKADNKKWSVTKCAERAMPGRLYVHPDSPAIGAHMROLVSFOKILITNNHLP			167
		126	PADNHYKFKCDCKMMVMYAGAEAPAMGRLYVHPDSPAIGAHMROLVSFOKILITNNHLP		185
QY	168	FGHIIINSNHNKQORPLHTLYKADENNGBSSKNTACTSNVFRRTALATSTQNHKTLOKI			227
		186	FGHIIINSNHNKQORPLHTLYKADENNAPFSKNTACTSNVFRRTSLVSTYONHKTLOKI		245
QY	228	ENNPFAKSGRGSDDMELHRMSPOSKERVAVPRSTVROKVASNN-----	SPRSSRA		280
		246	ENNPFAKSGRGSDDSL--RVARLOKKEPLVTSKSIIMOROLVSOLSKAPDVSPLHSHQIA		304
Db	281	LTSSMSGQYOCENG-----VSGPSODLLPRPNRYPLROHNSIYACTYKRGCECHPWS			335
QY	305	L-----OHUYENGANHQFAAERPOL--PLMTFPTQRODSLFEYGLCKLRDOSARHLDL			355
Db	336	ICFLSYL	342		
		356	PCRSTL	362	

RESULT 10

ID	09IAK9	PRELIMINARY:	PRT:	543 AA.
AC	09IAK9			
DT	01-OCT-2000 (TREMBLrel, 15, Created)			
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel, 20, Last annotation update)			
DE	T-box transcription factor tbx1.			
OS	Brachydanio rerio (Zedrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20130955; PubMed=10664151;			
RA	Ruvinsky I., Oates A.C., Silver L.M., Ho R.K.;			
RT	"The evolution of paired appendages in vertebrates: T-box genes in the			
RT	zebrafish."			
RL	Dev. Genes Evol. 210:82-91(2000).			
DR	EMBL; AF179406; AAF59836.1; -			
DR	HSSP; P24781; 1XBR.			
DR	InterPro; IPR001699; TF_T-box.			
DR	Pfam; PF00907; T-box; 1.			
DR	PRINTS; PR00937; TBOX.			
DR	SMART; SM00425; TBOX; 1.			
DR	PROSITE; PS01264; TBOX_2; 1.			
DR	PROSITE; PS50252; TBOX_3; 1.			
SO	SEQUENCE 543 AA; 60214 MW; AE2BBCA68BD3B9E7 CRC64;			

Query Match	59.88;	Score 1130.5;	DB 13;	Length 543;
Best Local Similarity	66.98;	Pred. No. 2.2e-94;		
Matches 230;	Conservative 32;	Mismatches 67;	Indels 15;	Gaps 7

```

QY 4ADGFGIATLPLEBDKADLDPCDSKESALGASKSPSSPOAALFOOGEGKIVYLHREL 63
   | | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 10 ADEGMITVAQSGGR---ELASDS----SHLGELT-TPSRPQNNEDQOSTENKIVYLHREL 62

QY 64 WLKFEHVEGEMITTKAGRMFPSTYKYVYGLNPKTKITLMDIYPADDDHRYKFDANKWSY 123
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 WKLIHAGEMITTKAGRMFPSTYKYVYGLNPKTKITLMDIYPADDDHRYKFDANKWVY 122

QY 124 TGKAEPMFGRLVHPDPSPATGAHMMROLVSFOKATLNNHLDPEGHITLMSMKYQPRL 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 AGKAEPAMFGRLVHPDPSPATGAHMMROLVSFOKATLNNHLDPEGHITLMSMKYQPRL 182

QY 184 HYYKAEDNNNGPFSKNTAFCTHYHFRETAFATYSYONHKTITOLKLENNPFAKFGSDOME 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 183 HYYKAEDNNNAFGSKNTAYCTHYHFHETAFSTVSYONHKTITOLKLENNPFAKFGSDOEG 242

QY 244 IHRMSDMQSKEXVPVPRSTROYKAVASNNSPFSESERA--LSTSSNLASQYCEMGVSGP- 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 243 L-RVSRLOCKDYPVLSKNNWRORLILSHGNHLSGKLSACVLASHPOVLSHYQYOGVYLPN 301

QY 301 --SQDILPPRPVPLQLEHNSOITYHCTKKKGEDDHWSICFLSYL 342
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 302 SDSQEAAL--SNSPTSSREPSLLYHCFKHHNDPRHLIELGCKRPYL 343

```

RESULT	11
042436	
ID	042436
AC	042436;
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	NTTbox1. (Fragment).
GN	NTTbox1.
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC	Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
NC	Notophthalmus.
OX	NCBI_TaxID=8316;
RN	[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=97236482; PubMed=9118806;
RA Simon H.G., Kittappa R., Khan P.A., Tsilfidis C., Liversage R.A.,
Oppenheimer S.;
RT "A novel family of T-box genes in urodele amphibian limb development
and regeneration: candidate genes involved in vertebrate
forelimb/hindlimb patterning";
RL Development 124:1355-1366(1997).
DR EMBL: U64433; AAB62743.1; -
DR HSSP: P24781; 1XR.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PFO0907; T-box; 1.
DR PRINTS: PRO0937; TBOX.
DR SMART: SM00425; TBOX; 1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
FT NON_TER 1
SQ SEQUENCE 411 AA; 46092 MW; 77FB2B5B007E675 CRC64;

Query Match 57.5%; Score 1085.5; DB 13; Length 411;
Best Local Similarity 87.0%; Pred. No. 1.9e-90;
Matches 207; Conservative 7; Mismatches 17; Indels 7; Gaps 4;

QY 108 PADDRHYFADNKKSVTGKAEAPAPGRLVYHPDSPATGAHMMROLVSFOKLTNNHLDLP 167
DB 1 PDDDRHYFADNKKSVTGKAEAPAPGRLVYHPDSPATGAHMMROLVSFOKLTNNHLDLP 60
QY 168 FGHILNSMHHYOPRLIIVKADNNGFGSKNTAFCTHVFPEFAFIAVTSYONHKITOLKI 227
DB 61 FGHILNSMHHYOPRLIIVKADNNGFGSKNTAFCTHVFSEFAFIAVTSYONHKITOLKI 120
QY 228 ENNPAPKGFRCSDMEHLHRMSRMSQKEYVYVPRSTVROKVASNPSFSSSRALSTSSNL 287
DB 121 ENNPAPKGFRCSDMEHLHRMSRMSQKEYVYVPRSTVROKVASNPSFSSSRALSTSSNL 179
QY 288 GQOYCENGVSQSDLLPPNPY-PLPOESHQIYHCKRKG--EC--DHPMSICFL 339
DB 180 SSQOYCENGVSSTQDILLPTTNPYOLNQHQIYHCKRKGDECSPTFHFKKSYM 237

RESULT 12
Q9IBC5 PRELIMINARY; PRT; 133 AA.
AC Q9IBC5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Transcription factor Tbx5 (Fragment).
GN Tbx5.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20171077; PubMed=10704879;
RA Takabatake Y., Takabatake T., Takeshima K.;
RT "Conserved and divergent expression of T-box genes Tbx2-Tbx5 in
Xenopus";
RL Mech. Dev. 91:433-437(2000).
DR EMBL: AB032944; BA03084.1; -
DR HSSP: P24781; 1XR.
DR TRANSFAC: T04423; -
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PFO0907; T-box; 1.
DR PRINTS: PRO0937; TBOX.
DR SMART: SM00425; TBOX; 1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
FT NON_TER 1
SQ SEQUENCE 133 AA; 15209 MW; 9F5A287F9400B4EA CRC64;

Query Match 37.7%; Score 712; DB 13; Length 133;
Best Local Similarity 98.5%; Pred. No. 4e-57;
Matches 131; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 118 DNKMSVTGKAPAPGRLVYHPDSPATGAHMMROLVSFOKLTNNHLDPEFGHILNSMH 177
DB 1 DNKMSVTGKAPAPGRLVYHPDSPATGAHMMROLVSFOKLTNNHLDPEFGHILNSMH 60
QY 178 KYOPRLHIVKADENNGFGSKNTAFCTHVFPEFAFIAVTSYONHKITOLKIENNPFAKGR 237
DB 61 KYOPRLHIVKADENNGFGSKNTAFCTHVFSEFAFIAVTSYONHKITOLKIENNPFAKGR 120
QY 238 GSDDMELHRMSRM 250
DB 121 GSDDMELHRMSRM 133

RESULT 13
Q9GOE7 PRELIMINARY; PRT; 174 AA.
AC Q9GOE7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE T-box protein Amphibx6/16 (Fragment).
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20519458; PubMed=11063699;
RA Ruvinsky I., Silver L.M., Gibson-Brown J.J.;
RT "Phylogenetic analysis of T-box genes demonstrates the importance of
amphioxus for understanding evolution of the vertebrate genome";
RL Genetics 156:1249-1257(2000).
DR EMBL: AF262565; AAC34890.1; -
DR HSSP: P24781; 1XR.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PFO0907; T-box; 1.
DR PRINTS: PRO0937; TBOX.
DR SMART: SM00425; TBOX; 1.
DR PROSITE: PS01283; TBOX_1; 1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
FT NON_TER 1
SQ SEQUENCE 174 AA; 20115 MW; EF36A4C31678E880 CRC64;

Query Match 35.7%; Score 674; DB 5; Length 174;
Best Local Similarity 70.3%; Pred. No. 1.7e-53;
Matches 123; Conservative 24; Mismatches 26; Indels 2; Gaps 1;

QY 63 LMLKFEVGTMITTKAGRRMFPSYKVTGILMPKTKTILMDIVPADDRHYFADNKKMS 122
DB 1 LMDSPHDIGTMITTKAGRRMFPTYKASISGIDPNAKYLILMDIVPADDRHYKHNSEW 60
QY 123 VTGKAEPLMPGRLIYHPDSPATGAHMMROLVSFOKLTNNHLDPEFGHILNSMHKYOFR 182
DB 61 VSGKAEPLMPGRLIYHPDSPATGCTOMMKOSVFIRKLKLTNNAMDOCGHILNSMHKYOFR 120
QY 183 LHIVKADENNGFGSKNTAFCTHVFPEFAFIAVTSYONHKITOLKIENNPFAKGR 237
DB 121 LHIVKADENNGFGSKNTAFCTHVFSEFAFIAVTSYONHKITOLKIENNPFAKGR 173

RESULT 14
Q9IBC6 PRELIMINARY; PRT; 132 AA.
AC Q9IBC6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DE		01-MAR-2002 (TREMblrel. 20, last annotation update)
DT		Transcription factor Tbx4 (Fragment).
GN		TBX4.
OS	Xenopus laevis (African clawed frog);	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;	
OC	Xenopodidae; Xenopus.	
OX	NCBI_TaxID=8355;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=20171077; PubMed=10704879;	
RA	Takabatake Y., Takabatake T., Takeshima K.;	
RT	"Conserved and divergent expression of T-box genes Tbx2-Tbx5 in Xenopus."	
RL	Mech. Dev. 91:433-437(2000).	
DR	EMBL; AB032943; BAA93083.1; -	
DR	HSSP; P24781; IXBK.	
DR	TRANSEAC; T04421; -	
DR	InterPro; IPRO01699; TF_T-box.	
DR	Pfam; PF00907; T-box; 1.	
DR	PRINTS; PR00937; TBOX.	
DR	SMART; SMO0425; TBOX_1.	
DR	PROSITE; PS01264; TBOX_2; 1.	
DR	PROSITE; PS02052; TBOX_3; 1.	
FT	NON_TER	1
ET		132
SO	SEQUENCE	132 AA; .14978 MW; EBC79ECBAAB83466 CRC64;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:44:00 ; Search time 39 Seconds
(Without alignments)
1192.422 Million cell updates/sec

Title: US-10-027-059A-1
Perfect score: 1889
Sequence: 1 MADADEGFIAGHTPLEPDAK.....CDHPWSICFLSYLFLSLGWS 349

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the chance being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	647.5	34.3	716	22	AAAM79757
2	643.5	34.1	722	22	AAAM78773
3	626	33.1	988	22	ABAB60007
4	624	33.0	295	22	ABAB5080
5	618	32.7	699	22	ABAB5542
6	595	31.5	360	22	ABAB1281
7	585	31.0	469	22	ABAB61330
8	562	29.8	424	22	ABAB61258
9	553.5	29.3	580	22	ABAB52524
10	541.5	28.7	660	22	ABAB2471

11	535.5	28.3	423	23	AAAB47994
12	535	28.3	530	22	AAAB51014
13	528.5	28.0	535	22	AAAB51013
14	517.5	27.4	365	21	AAAB57098
15	513	27.2	448	20	AAAB3773
16	507	26.8	686	21	AAAB57394
17	505	26.7	688	21	AAAB57393
18	499.5	26.4	517	21	AAAB49414
19	489	25.9	747	22	ABAB63136
20	207.5	11.0	105	20	AAAB3774
21	194.5	10.3	209	22	ABAB6193
22	179	9.5	156	22	AAAB25440
23	179	9.5	156	23	AAAB51291
24	177.5	9.4	938	22	AAAB94197
25	158	8.4	57	22	ABAB80026
26	158	8.4	57	22	ABAB33200
27	158	8.4	57	22	ABAB18665
28	158	8.4	57	22	AAAB53993
29	158	8.4	57	22	AAAB63383
30	158	8.4	57	22	AAAB14253
31	158	8.4	57	22	AAAB66653
32	158	8.4	57	22	AAAB01983
33	158	8.4	57	23	ABAB36032
34	129	6.8	90	22	AAAB04712
35	111	5.9	276	22	ABAB11017
36	102.5	5.4	872	22	ABAB7619
37	101	5.3	1047	22	ABAB67483
38	99	5.2	24	20	AAAB02871
39	98.5	5.2	422	17	AAAB7467
40	98.5	5.2	422	22	AAAB67901
41	98.5	5.2	422	22	AAAB67939
42	98.5	5.2	509	23	ABAB49686
43	98.5	5.2	971	22	ABAB63836
44	95.5	5.1	859	21	AAAB38841
45	95.5	5.1	886	21	AAAB38840

ALIGNMENTS

RESULT 1	
AAAM79757	AAAM79757 standard; Protein: 716 AA.
ID	AAAM79757
XX	AAAM79757
AC	AAAM79757
XX	06-NOV-2001 (first entry)
DT	
DE	Human protein SEQ ID NO 3403.
XX	
XX	Human: cytokine: cell proliferation: cell differentiation: gene therapy;
KW	vaccine: peptide therapy: stem cell growth factor: hematopoiesis;
KW	tissue growth factor: immunomodulatory; cancer: leukaemia;
KW	nervous system disorder: arthritis; inflammation.
OS	Homo sapiens.
XX	
XX	WO200157190-A2.
PN	
XX	09-AUG-2001.
PD	
XX	
XX	05-FEB-2001; 2001WO-0504098.
PF	
XX	
PR	03-FEB-2000; 2000US-0496914.
PR	27-APR-2000; 2000US-0560875.
PR	20-JUN-2000; 2000US-0598075.
PR	19-JUL-2000; 2000US-0620325.
PR	01-SEP-2000; 2000US-0654936.
PR	15-SEP-2000; 2000US-0663561.
PR	20-OCT-2000; 2000US-0693325.
PR	30-NOV-2000; 2000US-0728422.
XX	(HYSE-) HYSEQ INC.

zebrafish ntl. Br
murine T-bet prote
Human T-bet protei
Human prostate can
Human T-box polype
Human nerve differ
Mouse nerve differ
Human T-box transc
Drosophila melanog
Human T-box polype
Novel human diagno
Human mdtr protein
Human MDR1 SEQ ID
Human protein sequ
Human peptide #677
Peptide #706 encod
Protein #664 encod
Human brain expres
Human bone marrow
Peptide #687 encod
Peptide #670 encod
Peptide #665 encod
Human peptide enco
Human polypeptide
Human semaphorin 4
AMEPV mRNA caping
Drosophila melanog
Fragment of human
Glia growth facto
Human neuregulin 9
Human neuregulin 9
Listeria monocytog
Drosophila melanog
Arabidopsis thalia
Arabidopsis thalia

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 P1 Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 XX N-PSDB: AAK52890.
 PT
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 PS
 PS Claim 20: Page 324: 6221pp: English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAK78323-AAK80902) that exhibit activity elating to
 CC cytokine cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in either cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haemopoietic regulatory
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibit activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK60020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 XX Sequence 716 AA;

Query Match	Similarity	Score	E-value	DB	Length
Best Local	100.0%	647.5	0.0e+00	NCBI	716
Matches	145	Conservative	39	Mismatches	77
				Indels	41
				Gaps	7

Query	Subject	Score	E-value	DB	Length
5	DGGRIAMHFLIEDPDKLDCSDKSPSALGAPSKSPSSPMAF----	647.5	0.0e+00	NCBI	716
Db	57	DQLYGAAETGI-----PSSISLG-----	647.5	0.0e+00	NCBI
57	FLHEHEILMLKFNHEVGEEMITITAGRRMFSPYKYKATGMLPKTKYTLTLLMDIVPADNRHYK	647.5	0.0e+00	NCBI	116
97	HLAELEIMDDPFRKRGTEWITKSGRRMRPPRYKSGGLDKKAKYTLTLLMDIITAADDCRYK	647.5	0.0e+00	NCBI	156
117	ADNKSMTYGAERBAMDEGRULYVHPDSPAATGAMHMKOLVSFOKLTLLNNHLPFEGHIIINSM	647.5	0.0e+00	NCBI	176
157	NHSRRVWAGKADEMDPKRRMYIHPDSPAATGADQWMSKYVTFNKIKLTLLNNSIDKHGFTLLNSM	647.5	0.0e+00	NCBI	216
177	HKYOPRLHIYKADENNNGFSKNTYACTHVFPEPTAFIATVSYONHKITOLKIDNNPFAKGF	647.5	0.0e+00	NCBI	236
217	HKYOPREFIVRA--NDILKLPYSTERTLTPETEFIAIATVAYONDITOLKIDNNPFAKGF	647.5	0.0e+00	NCBI	274
237	R--GSDMDHEILHRMSRQS-----KEPVYPRSYRQKVAASHSPFSESEALSTLS	647.5	0.0e+00	NCBI	285
275	RDGTGGRKREKRKQTLTQSKRYVDERHNKKENGTSDSESSQDAFNCFAASSPAPRLKRTS	647.5	0.0e+00	NCBI	334
286	NL 287				
335	NL 336				

Result	Score	E-value	DB	Length
AA087773	647.5	0.0e+00	NCBI	716
AA087773	647.5	0.0e+00	NCBI	716
AA087773	647.5	0.0e+00	NCBI	716
06-NOV-2001	(first entry)			
Human protein seq ID	NO 1435			
Human: cytokine, cell proliferation, cell differentiation, gene therapy, vaccine, peptide therapy, stem cell growth factor, haematopoiesis, tissue growth factor, immunomodulatory, cancer, leukaemia, nervous system disorder, arthritis, inflammation.				

OS	Homo sapiens.
XX	
PN	W020015190-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US04098.
XX	
PR	03-FEB-2000; 2000US-0496914.
PR	27-APR-2000; 2000US-0560875.
PR	20-JUN-2000; 2000US-0598075.
PR	19-JUL-2000; 2000US-0620325.
PR	01-SEP-2000; 2000US-0654935.
PR	15-SEP-2000; 2000US-0663561.
PR	20-OCT-2000; 2000US-0693325.
PR	30-NOV-2000; 2000US-0728422.
XX	
PA	(HYSEQ) HYSEQ INC.
XX	
PI	Tang Y.T., Liu C., Drmanac R.T., Asundi V., Zhou P., Xu C., Cao Y., Zhang Q.Q., Wang D., Wang J., Zhang J., Ren F., Chen R., Wang Z.W.; Xue A.J., Yang Y., Wejhrman T., Goodrich R.
PI	
XX	
DR	WPI: 2001-476283/51.
DR	NFSDB: AAK51906.
XX	
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy -
XX	
XX	Claim 20; Page 3701-3702; 6221pp; English.

Query Match	Best Local Similarity	Matches	144: Conservative	43: Mismatches	90: Indels	33: Gaps	7: C	
5	DEGFGILHFFLEPDKADLP	CDSPESALGAPSSKSPSSQAA	----	TOGMEGI	----	KV	56	
66	DQILGAAGTEGI	-----	PFSSLG	-----	FOALRLPLKTMPEEEVEEDBP	KV	105	
57	FLHERLMLTAFHEVGETEMI	ITFKAGRNRFPSYKKVGL	YGLMLPRTKYIL	LMDDVPAD	DIHRYKF	116		
106	HLAEKELMDQFHKRGCEV	ITKSGRRMFPEPKV	RGSGDLKAKAYIL	LMDDIDDCRYKF	165			
117	ADNKWSYTGAEAPRPG	LYVHPDPSPTGSHM	ROLYSPOKLTK	LNHLLPFGHII	ILNSM	176		
156	HNSRMWACKADPEAK	PKMYIHPDSPTGOMSK	MYFHNKLT	KLNNISD	KHGFILNSM	225		
177	HKQYPLHLYKADENK	GGSKNTAFCHIVPE	FAFVATYV	SQNHKLTQ	LKIENNPAKGF	236		
226	HKQYPRRHLYRA	----	NDILKLPSTFK	RYLPEPEFLAV	ATAYATQNKIT	ITQLDIDNNPAKGF	283	
237	RGSDDMELHMRSG	WMOCKEYPVVPR	STVROKVA	SNHSPPSS	ESRALSTSSML	SGOYCENG	296	
284	RDTCNGRREKRKQ	LTLAGMRVDFDR	RHKKEG	NGTSDSE	-----	SSDQAFNCFQ	AOASSPAASTG	341

QY 297 VSGRPSODLDP 306
| : || |
Db 342 TSN-LKDLCP 350

RESULT 3
ABB60007
ID ABB60007 standard; Protein: 988 AA.

AC ABB60007;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 6813.

DE Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

DR N-PSDB: ABL04110.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

XX Disclosure: SEQ ID NO 6813; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
comprises genomic DNA sequences (ABU16176-ABU30511), expressed DNA
sequences (ABU1840-ABU16175) and the encoded proteins

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 988 AA:

Query Match 33.1%; Score 626; DB 22: Length 988;

Best Local Similarity 47.1%; Pred. No. 2e-52;

Matches 136; Conservative 36; Mismatches 69; Indels 46; Gaps 6;

QY 55 KVLHERELMLKHEVGTETMITKAGRRMPSPYKVTGMLPKRYLLMDIVADPHRY 114

Db 329 KVTLEGKDLWEKFKLTGEMVITKSGROMPFOMKFRVSGIDAKAKYLLLDIVAADRY 388

QY 115 KFDANKSVTGKAPAPGRLYVHPDSPAICAHMROLVSQKIKLTNNHLDPRGHT--- 171

Db 389 KFNHSRMVAVGKADPEMPKRYTHPDSPTGEGMOKVVSFHKLTNNISDKHGFSVT 448

QY 172 ILNSMHTYQPLTLVKADENGFGSKNTAFCTHVPETAFIAVSYONHKITOLKTEENP 231

Db 449 ILNSMHTYQPLTLVKADENGFGSKNTAFCTHVPETAFIAVSYONHKITOLKTEENP 506

QY 232 FAKGR--GSDDMELHRRMSQSKPYVPRSTVRKRVASNSPSSSRALSTSSNLGS 289

Db 507 FAKGRFDTGAGKRE-----KNCVRQALMSNR---GSDSKLTNTHTVSSS 547

QY 290 QYCEGVSGRPSODLDP-----PPNPYLPQHS 318
| : || |
Db 548 RAPLHLGHAGRRPHLHPHALLDNQDDDDKLDVGVGPSPPLPLSHS 596

RESULT 4

ABB95080
ID AAB95080 standard; Protein: 295 AA.

AC AAB95080;

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16968.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX Claim 8; SEQ ID 16968; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the combination
of the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.

XX Sequence 295 AA;

Query Match 33.0%; Score 624; DB 22; Length 295;
 Best Local Similarity 53.3%; Pred No. 5, 1e-53;
 Matches 122; Conservative 39; Mismatches 48; Indels 20; Gaps 6;

QY 11 AHTLEPDAKDLQCDKSPESALGAPSKSPSSFOAFTQOGMEGKVFLEHRELMLKFEHV 70
 DB 62 AHTPLPL-----LP-----PAMGT--EPAPSAPEAL---HSLPGVSLLENRELKMESSV 107
 QY 71 GTEMITRACGRMPSPYKVKVTGLNFKTKYILLMDIYPADDRKFAADNKWSYTGKAEP 130
 DB 108 GTEMITRACGRMPSPYKVKVTGLNFKTKYILLMDIYPADDRKFAADNKWSYTGKAEP 167
 QY 131 MFGRLVHPDSPAATGAMHMKQLVSEFOKLTNNHLDPEFGHILNSMKKYOPRLHIVKADE 190
 DB 168 LPDRYVTHPDSPATGAMHMKQPVSEFHRVYKLTNSLTDPHGHLILSHMKKYOPRLHIVRAAO 227
 QY 191 --NNGFGSKNTAFCTHVPETAFAIVTSYONKHTITOLKIENPFAGFR 237
 DB 228 LCSQHWG----GMAFPFETTFISYTAQNPQTLQKIANPFAKGR 272

RESULT 5
 ABB65542
 ID ABB65542 standard; Protein: 699 AA.
 AC ABB65542;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 23418.
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI: 2001-656860/75.
 DR N-PSDB: ABL09645.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Disclosure: SEQ ID NO 23418; 21pp + Sequence Listing: English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB101840-AB16175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WFO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC Sequence 699 AA;
 SO Query Match 32.7%; Score 618; DB 22; Length 699;

Best Local Similarity 42.6%; Pred No. 7, 4e-52;
 Matches 139; Conservative 42; Mismatches 95; Indels 50; Gaps 8;

QY 36 SKRSPSOAFAFTQOGMEGKVFLEHRELMLKFEHVGTETMITRACGRMPSPYKVKVTGLN 95
 DB 182 SKGSEFVHPDSPAATGAMHMKQLVSEFOKLTNNHLDPEFGHILNSMKKYOPRLHIVKADE 237
 QY 96 PKTKYILLMDIYPADDRKFAADNKWSYTGKAEPMPGRLVHPDSPAATGAMHMKQLV 133
 DB 238 PHATYICMDPVPMDKRYAFHNSCVYVAKADPISPRLHIVHPDSPAATGAMHMKQLV 297
 QY 154 SFQKLTNNHLDPEFGHILNSMKKYOPRLHIV-----KADENNGFGSKNTAFCTHVF 206
 DB 298 SFQKLTNNHLDPEFGHILNSMKKYOPRLHIVLPKKNASLDEN-----ENSHFRTFIF 353
 QY 207 PETAFIAVTSYONKHTITOLKIENPFAGFR--GSDDELHRSRMSQKEYPVYPR---- 260
 DB 354 PETSFYAVTAYQORVYOLKISSNPFAGFRDQDNDVTTGGSSKSSMSHESQARRKQ 413
 QY 261 ----STVRKVASNHSPPSSSRALSTSNIG-----SOYCENGV----- 297
 DB 414 QQQQQQQQQQQQQQQQQQQQQKERTAAATSNFGLSCELAIEQQQQQQQQQQQQGLDPATPS 473
 QY 298 ----SCPSODLPPNPYPYLPQESHQ 319
 DB 474 SSTSNSPDLQYQMEQQLQQQHQQ 499

RESULT 6
 ABB61281
 ID ABB61281 standard; Protein: 360 AA.
 AC ABB61281;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 10635.
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI: 2001-656860/75.
 DR N-PSDB: ABL05384.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Disclosure: SEQ ID NO 10635; 21pp + Sequence Listing: English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB101840-AB16175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WFO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC Sequence 360 AA;
 SO Query Match 32.7%; Score 618; DB 22; Length 360;

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 360 AA;

SO Query Match 31.5%; Score 595; DB 22; Length 360;

Best Local Similarity 45.3%; Pred. No. 5.1e-50;
Matches 124; Conservative 42; Mismatches 92; Indels 16; Gaps 5;

QY 39 PSSQAQAFQOQMGIVFLHERLMLKHEVGTETITKAGRMPPSYKVTGTLNPKT 98

Db 20 PSRPT-----LPGVAKLENDLMQOFKIGTETITKSGRMPPSMRVLISGLEEEA 72

QY 99 KYILMDIVPADHRYKFAADKMSVTKAPAMPGRILYHPDPSATGAHMMROLYSPOKL 158

Db 73 SYCVLEMLNPLGDCRYKFGSQWVPAGCAEPQSRMLHPDSPAATGAHMOQALNPKV 132

QY 159 KLTNNLDPFGHILLNSMHKYOPLRIIVKADENNGF-GSKNTAFCTHVPETAFIAVTSY 217

Db 133 KLTNNLDSGHIYVLASMHKYOPLRIIRSELTQLPMAPOQAF---VFPETEFVAVTAY 189

QY 218 QNHKITOLKIENNPFAGCFRGSDDMELHRMSRMOSKEYPVPRSTVKQKVASNHSFSSSE 277

Db 190 QNDKITLKDNNPFAGCFRGSQRCKR--KLNSGNSLTLESDGSSVSSCDSPQAKR 247

QY 278 SRALSTSSNLGSOYOCENGVSQPSODLLPPNPY 311

Db 248 QRODFEODSTGVSVPATYGAHP---VANPMSPT 278

QY 278 SRALSTSSNLGSOYOCENGVSQPSODLLPPNPY 311

Db 248 QRODFEODSTGVSVPATYGAHP---VANPMSPT 278

RESULT 7
ABB61330 ID ABB61330 standard; Protein: 469 AA.

AC ABB61330;
XX 26-MAR-2002 (first entry)

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 10782.

DE Drosophila melanogaster polypeptide SEQ ID NO 10782.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2001; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW.

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05433.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 10782; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins

CC (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 469 AA;

SO Query Match 31.0%; Score 585; DB 22; Length 469;

Best Local Similarity 43.5%; Pred. No. 7.4e-49;
Matches 130; Conservative 44; Mismatches 81; Indels 44; Gaps 8;

QY 51 MEGIKVFLHERLMLKHEVGTETITKAGRMPPSYKVTGTLNPKTYLLMDIYPAD 110

Db 53 LPGVEMTLQNDMDMKQFHQIGTETITKSGRMPPSMRVLISGLEDSNVCVLLLEWPIG 112

QY 111 DHRYPADNKMWSYVKAEPAMPGRILYHPDSPAATGAHMMROLYSPOKLKTNNLDPFGH 170

Db 113 DCRYKFGSQWVPAGCAEPQSRMLHPDSPAATGAHMOQALNPKVKTNNLDSGHI 172

QY 171 IILNSMHKYOPLRIIVK-ADENNGFSGSKNTAFCTHVPETAFIAVTSYQNHKITOLKIEN 229

Db 173 IVLASMHKYOPLRIIVRTADIAQIPMAPOQAF---VFAETEFVAVTAYQNDKITLKD 229

QY 230 NPFAKGR-----GSDMELHRMSRMOSK-EYRVPV-----RST 262

Db 230 NPFAKGR-----GSDMELHRMSRMOSK-EYRVPV-----RST 262

QY 263 VRQVNASNHSFSSSESRALSTSSNLGSOYOCENGVSQPSODLL--PPP-----NPYPL 313

Db 290 IGNSDGPQIKRLRNGSACSLSSSLDDQ-----SYPGASSLALGSPPHLHSHHPRL 343

RESULT 8
ABB61258 ID ABB61258 standard; Protein: 424 AA.

AC ABB61258;
XX 26-MAR-2002 (first entry)

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 10566.

DE Drosophila melanogaster polypeptide SEQ ID NO 10566.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2001; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW.

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05361.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 10566; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX
SQ Sequence 424 AA:
Query Match 29.8%; Score 562; DB 22; Length 424;
Best Local Similarity 37.8%; Pred. No. 1.2e-46;
Matches 133; Conservative 41; Mismatches 94; Indels 84; Gaps 10;
XX
OY 39 PSSQAATGQMGKIVFLHERELMKFHEVGTETITTKARRMFPYSKYVYTGILNPKT 98
Db PPPPPQPTLPG--DVEAKLENELMQOFHSTGTEITTKCGRRMFPMSRLSLGLEEA 103
OY 99 KYLLMDIYPADHRYKFPADNKNKSVTGKAPAMPGRRLYVHPDSPATGAHMKQLVSFOKL 158
Db 104 STCVLELVPIGDCRYKFGSGVWPAGAPSPQRMILHPESPATGKHMOQSALLFSKY 163
OY 159 KLTNNHLPFGHIIILNSMKHYOPRLHIVKADENNGF-GSKNTAFCTHYVETAFIAVTSY 217
Db 164 KLTNNLTLLNNGHIIYLAHMKYOPRLHIVIRTSIDGQIPMAPOQAF--IFPTEFIATVAY 220
OY 218 QNHKITQALKIENNPAPKFGSDMELHNSRMOSKEYRVVP-----RSTVROKVA-- 268
Db 221 QNDRTIKLKLNNPFAKFGESQSRCKR--KINDSPPTQPEANQVDRSPVSLAKRL 278
OY 269 -----SNH----- 271
Db 279 RLYEERPOHHHOSHSHVPMQRHYLLDALANFYVPAPPAIEYARHIGAVPSMAATVP 338
OY 272 -SPFG--SESRALSTSSNLGSOYOCEN-----GVSGPSQDLLPPNPYPLP 314
Db 339 PSFASVSSSSSGSTGSESAADREADADADTVDVVGTAAPAPPPSP-PVP 389
XX
XX
RESULT 9
AB862524
ID ABB62524 standard; Protein; 580 AA.
XX
AC ABB62524;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14364.
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL06627.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosures: SEQ ID NO 14364; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX
SQ Sequence 580 AA:
Query Match 29.3%; Score 553.5; DB 22; Length 580;
Best Local Similarity 37.4%; Pred. No. 1.3e-45;
Matches 125; Conservative 54; Mismatches 96; Indels 59; Gaps 10;
XX
OY 24 CDSKFESALGAPSKSPSPD-----AAFTGQMGKIVFLHERELMKFHEVGTETM 74
Db 147 CSTPQSTGCTEDSEERLTPEPVOKAKPIYSCNCDCLPVOCHELTEDRPHDGTETM 206
OY 75 IITKARRMFPYSKYVYTG--LNFKTYILLMDIYPADHRYKFPADN--WSYTKAE 128
Db 207 IITKGRMFPYVVSFSGPLROIOPADRYAVLMDIIPMSKRYRAYVHRSAMLVAGKAD 266
OY 129 PAMPGRILYHPDSPATGAHMKQLVSFOKLKLTNNHLPFGHIIILNSMKHYOPRLHIVYA 188
Db 267 PAPPRLYALHPDSFSEALRKQIVISFEKVKLTNNEMDKNQIYLNHMKYOPRLHIVRL 326
OY 189 DENNGFGS-----NNTAFCTHYVETAFIAVTSYONHKITQALKIENNPAPKFGS---- 239
Db 327 SHGQSTPNSPKELQDLHRTYVPEVETVAVTAAYONQITLKLIDSPPKAGCRDSSRLT 386
OY 240 ----DME--LHNSRMOSKEYRVVPFRSTVROKVAHNSIFSSSRALSTSSN----- 286
Db 387 DFDRPMFALLLEQQLSPRLRP---DPLMQCPAAGDP---SSWALFKAROHLMF 439
OY 287 -----LGSQYOCENCVSGPSQDLLPPNP 310
Db 440 GGNSPYQALMMPQMTQAAAAAGP-----PPPP 468
XX
XX
RESULT 10
AB862471
ID ABB62471 standard; Protein; 660 AA.
XX
AC ABB62471;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14205.
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX MPI: 2001-656860/75.
 DR N-PSDB: ABL06574.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 14205; 21pp + Sequence Listing: English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA
 CC sequences (AB57737-AB872072).
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 660 AA:
 SO

Query Match 28.7%; Score 541.5; DB 22; Length 660;
 Best Local Similarity 37.7%; Pred. No. 2.5e-44;
 Matches 126; Conservative 57; Mismatches 120; Indels 31; Gaps 10;

OY 9 GLATPLLPDAKDLPCDSKPSALGAPSKSPS---SPO--AAFTQOGMEGKIVLHEREL 63
 DB 229 GAAATPSPPPP--PSQSPEELERLSPESPAAQOPTIVASCNDDLLTPVOCHLETREL 286
 OY 64 WLKHEVETEMITKAGRMPPSTKYVKTG---LNPKTKYLLMDIYPADDRKRPD 119
 DB 287 WDKFHELETEMITKSGRRMPPYVSPSGPLROTPADRAVLLDVLPDSRRRYAYH 346
 OY 120 K--NSVTGKAPAMPGRLYVHPDSPATGAHMMROLVSOKLKTNNHLDPRGHITLNSMH 177
 DB 347 RSSNLYAGKADPPPSRITAYHPDCLSPALRKQYVSEKTKLTNNEDKSGQVYVLSMH 406
 OY 178 KYQPLRLIVKADENNGF--GS---KNTAFCTHVPETAFIATVSYNHRTIQLKIENPF 232
 DB 407 KYQPLRLIVLRLSHGSIIPSGPELODMHKTFVPEYFAVTAQNLITRLKLTDSNPF 466
 OY 223 AKGFGSGDMELHNSRWQS---KEYVYPRSTYRKQVVASNHSFSSSRALSTSSNLG 288
 DB 467 AKGFGSSRLSDFDPRDAEFEEFDHMTAPLRFPPDPLMSQLTPOEADASMLLEKAR 526
 OY 289 SOYOCENGVSQSDLLP-----PPNPYPLP 314
 DB 527 QHLOW-FGRSPYTEMLPRLYKRSAAAPPPPPAP 559

RESULT 11
 AAB47994
 ID AAB47994 standard; Protein: 423 AA.

XX AAB47994;
 XX
 XX 22-JUL-2002 (first entry)
 XX
 XX zebrafish ntl.
 XX
 XX Promoter; ribozyme; RNA polymerase; zebrafish; morphological;
 KM behavioural; change; light tropism; chemical tropism; identification;
 KM disease related gene; ntl; mouse; brachyury; Xbra; Xenopus.
 XX
 OS Brachydanio rerio.
 XX
 XX US6355415-B1.
 XX

PD 12-MAR-2002.
 XX
 XX 29-SEP-1997; 97US-0939366.
 PF
 XX 29-SEP-1997; 97US-0939366.
 PR
 XX (UYOH-) UNIV OHIO.
 PA
 XX Wagner TE, Xie Y;
 PI
 XX MPI: 2002-391647/42.
 DR
 XX N-PSDB: AA172809.
 DR
 XX
 PT Detecting changes caused by cleaving mRNA encoded by gene of interest
 PT in zebrafish, by introducing ribozyme that cleaves RNA sequence into
 PT zebrafish cell having RNA encoding sequence, and detecting changes in
 PT cell -
 XX
 PS Example 1; Fig 1; 54pp; English.

CC This sequence is encoded by the zebrafish ntl gene, which is a homologue
 CC of the mouse *n* (brachyury) gene and the *Xbra* gene of *Xenopus*. This
 CC sequence was used to demonstrate the method of the invention as its
 CC sequence, and its function were known. The method of the invention is
 CC for detecting changes caused by cleaving mRNA encoded by gene of
 CC interest in zebrafish, a zebrafish cell or tissue, and involves
 CC providing the cell having a sequence homologous to a human sequence,
 CC where the sequence encodes an RNA sequence containing a substrate
 CC cleavage sequence, and a ribozyme sequence capable of cleaving the
 CC RNA substrate, introducing the ribozyme into the cell to generate a
 CC manipulated cell, and detecting changes in the manipulated cell. The
 CC ribozyme sequence is introduced into the zebrafish cell under conditions
 CC such that the RNA sequence is cleaved by the ribozyme sequence. The
 CC changes which are detected are selected from morphological changes and
 CC behavioural changes such as light tropism or chemical tropism in the
 CC manipulated cell relative to the zebrafish cell, or in tissue obtained
 CC from the manipulated cell relative to tissue obtained from the
 CC zebrafish cell. The method is useful for identifying one or more
 CC functions of a nucleic acid sequence in an organism which are useful
 CC for the rapid identification of, e.g. disease related genes which may
 CC be targeted for the treatment or prevention of disease.
 CC
 XX Sequence 423 AA:
 SO

Query Match 28.3%; Score 535.5; DB 23; Length 423;
 Best Local Similarity 41.3%; Pred. No. 3e-44;
 Matches 124; Conservative 40; Mismatches 93; Indels 43; Gaps 9;

OY 40 SSPQAAFTQOCME-----GIKVFLLHERELMLKFEVGTETMITKAGRMPPSYKRYT 92
 DB 15 SAVESEF-QKSGEKDASERDCLKLEDAELMTKRELTNEMITKGRMRPPVLRASVT 73
 OY 93 GLNPKTKYTLMDIYPADDRYKFEADNRKWSYTKAEPMPCRLVYHPDSPATGAHMMROL 152
 DB 74 GLDPNAMYSVLLDFVAADNNRKYVCGEVPQKPEPSCVYTHPDSNPGAHMMKAP 133
 OY 153 VSFQKLTNTNLDLDPFGHIIINSMHKYQRLIYVADENNGSGKNTAFCTIVPEFTAFI 212
 DB 134 VSFQKLVKLSNK-LNCGGQIMLSLHKYEPRIHIVV-----GTIOKMISSQSPFETFT 186
 OY 213 AVTSYNHRTIQLKTENNPFAGFGSDMDLHNSRWQSKEYPVVPSSTVYRKQVANSNS 272
 DB 187 AVTAQONEITALKTKHNPFAKAF-----LDAKERDRHKEVP-----DHS 226
 OY 273 PPSSESRALSTSSNLGSOYOCENGVSQSDLLPPNPYPLPQEHSQLYHCTKRKGECDH 332
 DB 227 TDNQO-----SGYSOLGWFPLPSNCGMPGSSSS---PPQIGAP-VHSSGSCYERYSSLRNH 278

RESULT 12
 AAB51014
 ID AAB51014 standard; Protein: 530 AA.

XX

AC AAB51014;
XX 20-MAR-2001 (first entry)
DT
DE Murine T-bet protein.
XX
XX Murine: T-bet; T box expressed in T cells protein; autoimmune disease;
KW Th2-type dysfunction.
XX
XX Mus musculus.
XX
XX MO200073453-A1.
XX
XX 07-DEC-2000.
XX
XX 01-JUN-2000; 2000MO-US15345.
XX
XX 02-JUN-1999; 99US-0137085.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Glimcher LH, Szabo SJ;
XX
XX WPI: 2001-061537/07.
XX
XX N-PSDB: AAC91989.
XX
XX Novel protein, T-bet, a new member of T-box family of transcription
PT factors is useful for identifying modulators of T-bet activity which
PT treat allergies, cancers, autoimmune and infectious diseases .
XX
XX Claim 17; Page 123-124; 126pp; English.
XX
XX The present sequence is murine T box expressed in T cells (T-bet)
CC protein. T-bet can be used in the treatment of autoimmune diseases
CC associated with a Th2-type dysfunction e.g. diabetes mellitus, arthritis,
CC multiple sclerosis, autoimmune thyroiditis, psoriasis, multiple
CC keratoconjunctivitis, alopecia areata, allergic responses due to
CC arthropod bite reactions, Crohn's disease, aphthous ulcer, iritis,
CC conjunctivitis, keratoconjunctivitis, ulcerative colitis, asthma,
CC allergic asthma, cutaneous lupus, erythematous, scleroderma, vaginitis,
CC proctitis, and drug eruptions. The role of T-bet is to drive
CC differentiation of Th cells.
XX
XX Sequence 530 AA;
SQ
Query Match 28.3%; Score 535; DB 22; Length 530;
Best Local Similarity 35.2%; Pred. No. 8e-44;
Matches 147; Conservative 40; Mismatches 97; Indels 134; Gaps 16;
OY 16 EPDAKDLPCSKPESALGAP-----SKSPSPQAAF-----TQO 49
DB 38 EFGADP-PTDRRAGSSLGTFYSGALVPARGRLGSAFAPRAQVAGPFGPGEFPFPA 96
OY 50 GMEG-----IKVFLHERELMKFHEVGTGM 74
DB 97 GAEYFVVDGVPADPRAGLYPGPREDYALPAGLEVSGKRLRALNHLMSKFNQHTGM 156
OY 75 IITAGRRMPSPYKVVYTGILNPKTKYILMDIVPADDRHKYKFNKSVYTGAEAPAMG- 133
DB 157 IITGGRMRPPLSTFVAGLEPTSHYRMFVDVVLVDQHHWRYSQKWCCKAGDSMCGN 216
OY 134 RLYVHPSPATGAHMMKQLVSFOKLTNN--HLDPEGHIILNSMKRQPLRIHYKAD 190
DB 217 RLYVHPSPATGAHMMKQLVSFOKLTNN--HLDPEGHIILNSMKRQPLRIHYKAD 276
OY 191 NNGGSKMTAFV-----THV--FPETAFIAYTSYONHKTQTKLENNPFKFGFR----- 237
DB 277 -----GPEFAACASNTHTVFTFOETOFIAYAYONAEITOLKIDNNPFAKGFRENFESMY 331
OY 238 -----GSDMLHMRMSRQSKKEYPV-----VPRSTVROK--- 266
DB 332 ASVDTSVSPGPGNCOILGDDP-----FSPILSNQVYPSRFPDLPQCPKMDISQPYWL 386

OY 267 -VASNHSPFSESRAIISTSNLGSYOCENGVSGPS-----ODLLPPNPYPLEPOH 317
DB 387 GTPREHS-YAEAFRAVSMKPTL-----LPSAPGPVTPYRYGQDVLAPGAWPAPOV 437
RESULT 13
AAB51013
ID AAB51013 standard; protein; 535 AA.
XX
XX AAB51013;
XX
XX 20-MAR-2001 (first entry)
XX
XX Human T-bet protein.
XX
XX Human: T-bet; T box expressed in T cells protein; autoimmune disease;
KW Th2-type dysfunction.
XX
XX Homo sapiens.
XX
XX MO200073453-A1.
XX
XX 07-DEC-2000.
XX
XX 01-JUN-2000; 2000MO-US15345.
XX
XX 02-JUN-1999; 99US-0137085.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Glimcher LH, Szabo SJ;
XX
XX WPI: 2001-061537/07.
XX
XX N-PSDB: AAC91989.
XX
XX Novel protein, T-bet, a new member of T-box family of transcription
PT factors is useful for identifying modulators of T-bet activity which
PT treat allergies, cancers, autoimmune and infectious diseases .
XX
XX Claim 16; Page 118-120; 126pp; English.
XX
XX The present sequence is human T box expressed in T cells (T-bet) protein.
CC T-bet can be used in the treatment of autoimmune diseases associated with
CC a Th2-type dysfunction e.g. diabetes mellitus, arthritis, multiple
CC sclerosis, autoimmune thyroiditis, psoriasis, keratoconjunctivitis,
CC alopecia areata, allergic responses due to arthropod bite reactions,
CC Crohn's disease, aphthous ulcer, iritis, conjunctivitis,
CC keratoconjunctivitis, ulcerative colitis, asthma, allergic asthma,
CC cutaneous lupus, erythematous, scleroderma, vaginitis, proctitis, and
CC drug eruptions. The role of T-bet is to drive differentiation of Th
CC cells.
XX
XX Sequence 535 AA;
SQ
Query Match 28.0%; Score 528.5; DB 22; Length 535;
Best Local Similarity 37.7%; Pred. No. 3.5e-43;
Matches 140; Conservative 41; Mismatches 115; Indels 75; Gaps 12;
OY 3 DADG--GEGLATPPEPAKDLPCSKPESALGAPSKSPSPQAAFQOQ-----ME 52
DB 44 DADERRGGGSLGSPY--FGALVP--APPSRFLCAVAYPRROAGFPQADESPPADAE 100
OY 53 G-----IKVFLHERELMKFHEVGTGM 77
DB 101 GYQPGEGYAAPDPRAGLYPGPREDYALPAGLEVSGKRLRALNHLMSKFNQHTM 160
OY 78 KACRRMPSPYKVVYTGILNPKTKYILMDIVPADDRHKYKFNKSVYTGAEAPAMG-RLY 136
DB 161 KQGRMRPPLSTFVAGLEPTSHYRMFVDVVLVDQHHWRYSQKWCCKAGDSMCGNRLY 220
OY 137 VHDSPATGAHMMKQLVSFOKLTNN--HLDPEGHIILNSMKRQPLRIHYKADENN 193
DB 221 VHDSPATGAHMMKQLVSFOKLTNN--HLDPEGHIILNSMKRQPLRIHYKADENN 280

```

0Y      194 FGSKNPAFCTHV--PPEFAFIATVSYQNHKTKTOLKIENNPAPAKFGSDMDLHRMSRMQ 251
           | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db       281 EAACN-ASNMHIFTEOTGFIAVTAYQNAEIQLIDNNPFAPGRREN---PESMTSV 339
OY      252 SKEPVPVRSTYOKWKAASHSPSSSRALSTSNIAGSOYCEN----GVSGPSODLLPP 307
           | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db       336 DTIPSPRPGRNCQFLGGDHISPL-----LPNQIVFPVRFHPDLPGAKADVPO 383
OY      308 PNDYPLEQHS 318
           ||:||
Db       384 AYWLGAPRDHS 394

RESULT 14
AAB57098
XX ID AAB57098 standard; Protein; 365 AA.
XX AC
XX AB AAB57098;
XX DT 13-MAR-2001 (first entry)
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1676.
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnerrary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
WM wound; infectious disease.
XX
OS Homo sapiens.
XX MO W020005174-AI.
PN PD 21-SEP-2000.
XX PE * -08-MAR-2000; 2000MC-US05988.
FR PR 12-MAR-1999; 99US-0124270.
XX PA ((HUMA-) HUMAN GENOME SCI INC.
XX PP (ROSE/) ROSEN C A.
PI Rosen CA, Ruben SM;-
XX WP1: 2000-587513/55.
DR N-PDSB; AAF16301.
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
Claim 11: Page 2144-2145; 233bp; English.
XX
AAFI5566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56365 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
XX invention.
XX
Sequence 365 AA;

```

[illegible]

RESULT 15
ID AAY03773 standard; Protein: 448 AA.
XX
AC AAY03773;
XX
DT 11-JUN-1999 (first entry)
XX
DE Human T-Box polypeptide, Tbx19.
XX
KW T-Box polypeptide; Tbx19; Holt-Oram syndrome; Ulnar Mammary syndrome; DiGeorge syndrome; cancer; psychiatric disorder; heart defect; human.
XX
OS Homo sapiens.
OS
PN WO911783-A2.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 96MO-EP05713.
XX
PR 17-AUG-1998; 98GB-0017900.
PR 03-SEP-1997; 97EP-0306830.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC. (UYNO-) UNIV NOTTINGHAM.
XX
PI Brook JD, Terrett JA, Yi CH;
XX
DR WPT, 1999-205187/17.
DR N-PSDB: AAX31841.
XX
PT Tbx19 protein - used to treat Holt-Oram syndrome, Ulnar Mammary syndrome, DiGeorge syndrome, cancer, psychiatric disorders, and heart defects
XX
PS Claim 3; Page 23; 32pp; English.
XX
CC This represents a T-box polypeptide having Tbx19 activity. Host cells containing a vector comprising the Tbx19 nucleic acids are used for the recombinant expression of the protein. The Tbx19 polypeptide can be used for treating Holt-Oram syndrome, Ulnar Mammary syndrome, DiGeorge syndrome, cancer, psychiatric disorders, and heart defects.
XX
SO Sequence 448 AA;
Query Match 27.2%; Score 513; DB 20; Length 448; Best Local Similarity 39.2%; Pred. No. 9,3e-42; Matches 113; Conservative 43; Mismatches 100; Indels 32; Gaps 5;

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rag

Page 10

Qy 54 IKVFLHERELMLKPEHVGTEMIITKAGRRMPSPYKVKYTGILNPKTKYLLMDIVPADDR 113
Db 41 LQIILEDAPLMOREFEVNEMIVTKNGRRMPVLKISYTGIDPNAMYSLLDVFPDTSR 100
Qy 114 YKPADNKSVTGKAPAMPGRLYVHPDSPATGAHMMROLVSFOKIKLTNNHLDPGHIL 173
Db 101 WKYVNGEWPVPAKPEVSSHSCVYIHPDSPWFGAHMMKAPISFSKYKLTNK-LNGCGOIML 159
Qy 174 NSMHKYOPRLHIVKADENNGFGSKNTAFCTHVPETAFAIVTSYONHKITOLKIENNPA 233
Db 160 NSLHKYEPOVHIVRY-----GSAHRKVTNCSFPETOFIATVAQNEETALKIKYNPFA 213
Qy 234 KGFSGDDMLHRMSRMSQKEYPVVPRSTVROKVASNHSFSSSRALSTSSNIGSOYOC 293
Db 214 KAF-----LDAKERNLRDVPEA-----ISESQHYTYSHLGWMIFSN 250
Qy 294 ENGV--SGPSODLLPPNPYPLPOEHSQIYHCTKRKGECDHMSICFL 339
Db 251 PDGVCTAGNSNQYAPLPLDPAFTHGCEHYSGLNGHRQAPYPSAYM 298

Search completed: June 17, 2003, 09:49:01
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:47:23 ; Search time 14 Seconds
(without alignments)
733.471 Million cell updates/sec

Title: US-10-027-059A-1
Perfect score: 1889
Sequence: 1 MADADGFGIAHTPLEPDAK.....CDHPMSICFLSLFLSLGWS 349

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	535.5	28.3	423	US-08-939-366-6	Sequence 6, Appl1
2	510	27.0	567	US-09-188-811-2	Sequence 2, Appl1
3	506.5	26.8	517	US-09-189-760-2	Sequence 2, Appl1
4	506.5	26.8	517	US-09-514-422-2	Sequence 2, Appl1
5	99	5.2	24	US-09-227-357-522	Sequence 522, App
6	98.5	5.2	411	US-08-470-339-189	Sequence 189, App
7	98.5	5.2	422	US-08-467-602-324	Sequence 324, App
8	98.5	5.2	456	US-08-467-602-366	Sequence 366, App
9	98.5	5.2	601	US-08-470-335-233	Sequence 233, App
10	98.5	5.2	601	US-08-467-602-323	Sequence 323, App
11	98.5	5.2	610	US-08-470-335-236	Sequence 236, App
12	98.5	5.2	610	US-08-467-602-332	Sequence 332, App
13	98.5	5.2	635	US-08-467-602-365	Sequence 365, App
14	98.5	5.2	644	US-08-467-602-374	Sequence 374, App
15	98.5	5.2	818	US-08-470-335-234	Sequence 234, App
16	98.5	5.2	818	US-08-467-602-321	Sequence 321, App
17	98.5	5.2	827	US-08-470-335-237	Sequence 237, App
18	98.5	5.2	827	US-08-467-602-333	Sequence 333, App
19	98.5	5.2	852	US-08-467-602-363	Sequence 363, App
20	98.5	5.2	861	US-08-467-602-375	Sequence 375, App
21	98.5	5.2	865	US-08-470-335-235	Sequence 235, App
22	98.5	5.2	865	US-08-467-602-322	Sequence 322, App
23	98.5	5.2	874	US-08-470-335-238	Sequence 238, App
24	98.5	5.2	874	US-08-467-602-334	Sequence 334, App
25	98.5	5.2	899	US-08-467-602-364	Sequence 364, App
26	98.5	5.2	908	US-08-467-602-376	Sequence 376, App
27	94.5	5.0	422	US-08-036-555B-170	Sequence 170, App

28	94.5	5.0	422	1	US-08-469-569-170	Sequence 170, App
29	94.5	5.0	422	1	US-08-428-926-3	Sequence 3, Appl1
30	94.5	5.0	422	1	US-08-249-322A-170	Sequence 170, App
31	94.5	5.0	422	1	US-08-428-927-3	Sequence 3, Appl1
32	94.5	5.0	422	1	US-08-428-298-3	Sequence 3, Appl1
33	94.5	5.0	422	1	US-08-339-517-3	Sequence 3, Appl1
34	94.5	5.0	422	1	US-08-469-526A-170	Sequence 170, App
35	94.5	5.0	422	2	US-08-734-591A-170	Sequence 170, App
36	94.5	5.0	422	2	US-08-469-660-170	Sequence 170, App
37	94.5	5.0	422	3	US-08-341-018-72	Sequence 72, Appl1
38	94.5	5.0	422	4	US-08-470-335-170	Sequence 170, App
39	94.5	5.0	422	4	US-08-735-021-170	Sequence 170, App
40	94.5	5.0	422	4	US-08-734-664A-170	Sequence 170, App
41	94.5	5.0	422	4	US-08-470-339-170	Sequence 170, App
42	94.5	5.0	422	4	US-08-467-602-170	Sequence 170, App
43	94.5	5.0	422	5	PCT-US94-05083C-166	Sequence 166, App
44	94.5	5.0	422	5	PCT-US94-05083C-185	Sequence 185, App
45	94.5	5.0	422	5	PCT-US95-06846A-170	Sequence 170, App

ALIGNMENTS

```
RESULT 1
US-08-939-366-6
: Sequence 6, Application US/08939366
: Patent No. 6355415
:
: GENERAL INFORMATION:
: APPLICANT: Wagner, Thomas E.
: TITLE OF INVENTION: Compositions and Methods for the Use of
: TITLE OF INVENTION: Ribozymes to Determine Gene Function
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESSES:
: ADDRESSER: Medien & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/939,366
: FILING DATE: 29-SEP-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Carroll, Peter G.
: REGISTRATION NUMBER: 32,637
: REFERENCE/DOCKET NUMBER: OHU-02749
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 423 amino acids
: TYPE: amino acids
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-939-366-6
:
: Query Match      28.3%; Score 535.5; DB 4; Length 423;
: Best Local Similarity 41.3%; Pred. No. 3.6e-50;
: Matches 124; Conservative 40; Mismatches 93; Indels 43; Gaps 9;
:
: QY 40 SSPQAAFGQGGME-----GKIVFLHREHRLMKHEVGTMLITKAGRMPPSYKVVKT 92
:   |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: DB 15 SAVSESEF-QKSGEKDASERDIKLSLEDEMLTKRKELNENIVTKTGRRMPVLRASVT 73
:   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: QY 93 GINPKTKYILLMDIVPADDRYKFAADNKWSVTGKAEPAMFGRLVYVHPDSPATGAHMMROL 152
```

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rai

Page 2

Db	74	GIDPAMSTYLLDEPVAAIDNNRMKRYVNGEYVGGKREPOQSSCYIHPDSDPNEGAIIMMKAP	133
Oy	153	VSFOKLTJTNHLDPEGHIIILSMHKYOPRLHIYKADENNGFGSKMTAFCTHVPETAFI	212
Db	134	VSFSKVTAKSNK-LNGGQITLMSLKHRYERHIVK-----GGIOKMSSSPPEQOFI	186
Oy	213	ATVSYQNHKTITQKIENNPPAFGFGSDOMELHMSRMOCKEYPVYPNSTVROKYASNHS	272
Db	187	ATYATQNEITLALIKINPPAKAF-----LDKESDREVP-----DHS	226
Oy	273	PFSSSRALSTSNLSQYOCERQYSGFSQDLLPVPNYPFLPOHISQIYHCTKRKGECDH	332
Db	227	TDNQO----SYISDLGWFPLSNGMGSSS---PFOITAP-VHSSGICIRISLSLNH	278

RESULT 2
 US-09-188-811-2
 Sequence 2, Application us/09188811
 Patent No. 6037148
 GENERAL INFORMATION:
 APPLICANT: Khodadoust, Mehran
 TITLE OF INVENTION: NOVEL MTPX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 TITLE OF INVENTION: THEREFOR
 FILE REFERENCE: NMI-046CP
 CURRENT APPLICATION NUMBER: US/09/188,811
 CURRENT FILING DATE: 1998-11-09
 EARLIER APPLICATION NUMBER: 09/163,116
 EARLIER FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 567
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-188-811-2

Query Match	27.0%	Score 510:	DB 3:	Length 567:
Best Local Similarity	35.0%	Pred. No. 3,7e-47:		
Matches 139:	Conservative 50:	Mismatches 112:	Indels 96:	Gaps 14
Qy	30	SALAPAKSP-----	SSFOAPFOOCHET-----	KYFLHERELMK 66
Db	82	AAAGAAARAAIKYKKGCSALRARTLLEPAAVSCGGIGGVGSGFAHAYLILNRPMLK 141		
Qy	67	FHEVGETEMITTKAGRMFPSTYKVKVGLNPKRTYILMDIYPADHRYKFAADNKKMSVTK 126		
Db	142	FHRHQTETMITTKGGRMRFPPLSFNINLNPATAHYVFEVVLADENHMRFOCGKAWYTCCK 201		
Qy	127	AEPAMP-RLYVHPDSPATGAHMRKOLVSPQTKLNLN---ILDFEGHILINSMHKYOR 182		
Db	202	ADNNQGNKKYVHPESPNTGSHMRKQELSPCKLKLNNKGANNNTQVILQSLHKTQYR 261		
Qy	163	LHIYKADENN---GRGSAINTACTHYVPETALIVTSYONHKITOLKLENNFPKAGFPG 238		
Db	262	LHIVEETEDVEDLNEPSKYQTF---TFSETOELIATAQNTDITQKIDHNFPAAGFRD 318		
Qy	239	SDD-----MELHRMSHOS---KEYPVVPSTYRQKVAS-----NHSP---FSSSESR 280		
Db	319	NYDSMYTASENDRLTSPDPSRSHQIIVPGG--RYCVOSFPPEPVMTLPQARYNGERT 376		
Qy	281	LSTSNIGSIOYCCENGKVGSGPSODLLPPPN-----PY-----PL 313		
Db	377	VPOINGILISQOSEEVANPQRMILVTVYVQOQPNKILDISYSESYTSTLLPVGISLPL 436		
Qy	314	POHSQIYH-----CTRKCECDHWS 335		
Db	437	QTSHALGYPDPTFPAMMGCGRGSYRKMAAGLPLMT 473		

RESULT 3
US-09-189-760-2
; Sequence 2, Application US/09189760

```

? Patent No. 6031078
? GENERAL INFORMATION:
? APPLICANT: Rhodadoust, Mehran
? TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
? TITLE OF INVENTION: THEREFOR
? FILE REFERENCE: MM1-046CP2
? CURRENT APPLICATION NUMBER: US/09/189,760
? CURRENT FILING DATE: 1998-11-10
? EARLIER APPLICATION NUMBER: 09/163,116
? EARLIER FILING DATE: 1998-09-29
? EARLIER APPLICATION NUMBER: 60/089,467
? EARLIER FILING DATE: 1998-06-16
? EARLIER APPLICATION NUMBER: (PENDING)
? EARLIER FILING DATE: 1998-11-09
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO. 2
? LENGTH: 517
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-189-760-2

```

[illegible]

RESULT 4
 US-09-514-422-2
 : Sequence 2, Application US/09514422
 : Patent No. 6291193
 : GENERAL INFORMATION:
 : APPLICANT: Rhododoust, Mehran
 : TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 : TITLE OF INVENTION: THEREFOR
 : FILE REFERENCE: MNI-046CP2
 : CURRENT APPLICATION NUMBER: US/09/514,422
 : CURRENT FILING DATE: 2000-02-28
 : PRIOR APPLICATION NUMBER: US/09/189,760
 : PRIOR FILING DATE: 1998-11-10
 : PRIOR APPLICATION NUMBER: 09/163,116
 : PRIOR FILING DATE: 1998-09-29
 : PRIOR APPLICATION NUMBER: 60/089,467
 : PRIOR FILING DATE: 1998-06-16
 : PRIOR APPLICATION NUMBER: (PENDING)


```

Db 202 -----WAVAGL- KRDSLLTVL-----GTW-----GHPAF 227
QY 132 P--GR-----YHDPSPATGAMHRKOLYSOKLKLTHNNHLDPPGHILLMSMAYOP 181
    |||
    |||
Db 228 PSCGRLEKEDSRIFYEMEPEDNSTSAPAPAFRASPPLE-TGRNLKEVSRYLCKRCALP 286
    |||
    |||
QY 182 RLHTVKADENNEGSKNTAFCTHVEPEAFIAVNSYON-----HKITOKTIENNFEAK 234
    |||
    |||
Db 287 RLKEKMSQE-SAAGSKLYLRC-ETISEVSSLRFKFKAGNLANLKNKPONIKLOKRP-- 341
    |||
    |||
QY 235 GFRGSDMELAKRSMQSEKEYPVVERSTYROKVASNHSFSSSRALSTSNLSOYO-- 292
    |||
    |||
Db 342 ---GKSELIRINKASLADSGEYWCCKVYSIKLGNDSASANITI-VESNATSTST-CTSHLYK 396
    |||
    |||
QY 293 -----CENG-----VSGPQDILLPPPNRY 311
    |||
    |||
Db 397 CAEKERTFCVNGGECPMWKDLSPNRYLCKCENEE 431

```

```

RESULT 9
US-08-470-335-233
: Sequence 223, Application US/08470335F
: Patent No. 6147190
: GENERAL INFORMATION:
: APPLICANT: GOODBART, ANDREW
: APPLICANT: STROOBANT, PAUL
: APPLICANT: MINGHETTI, LUISA
: APPLICANT: WATERFIELD, MICHAEL
: APPLICANT: MARCHIONNI, MARK
: APPLICANT: CHEN, MARIO S.
: APPLICANT: HILES, IAN
: TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
: FILE OF INVENTION: PREPARATION AND USE
: FILE REFERENCE: 04585/00200B
: CURRENT APPLICATION NUMBER: US/08/470,335F
: CURRENT FILING DATE: 1995-06-06
: EARLIER APPLICATION NUMBER: 08/036,555
: EARLIER FILING DATE: 1993-03-24
: NUMBER OF SEQ ID NOS: 252
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 233
: LENGTH: 601
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (601)...(601)
: OTHER INFORMATION: xaa is Arg.
US-08-470-335-233

```

Query Match: 5.2%; Score 98.5; DB 4; Length 601;
 Best Local Similarity 23.0%; Pred. No. 0.07; Indels 99;
 Matches 77; Conservative 40; Mismatches 119; Gaps 18;

```

QY 25 DSKPES-----ALGAPERS-----PSSPOAAFTQOGMGKIVFLHERETLYKPEYVG 71
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 114 DREPPAAGRALCPRAPEERLLAANGVBSMPAPRPSGEG-----PEAPYLKQYV- 167
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 72 TEKITTKACRRMPSTKVVAVGTGNPKTKYILIMDIYPADDHRYKFPADKNKWSVTKAEPRM 131
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 -----WAVYAGGL--KKDLSLLTVRL-----GTW-----GHPAF 193
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 132 P--GRL-----YVDPSPATGAAHMRKQLVSPCKLTNNHLDRECHITLSMHHYOP 183
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 PSCGRLEKEDSRITYFEHPEPANSRARAFAFRASFPLE--TGRLNKKKEYSVLCKRCALDP 255
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 RLHIYKADENNGGSKNTAFCTHVEPETAFLAVTSYON-----HKITQKLENNPFAK 234
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 RLKEMKSOE--SAAGSKYLTVRC--ETSEVYSILFFKFKFNGNELNKRKKPQNIKIQKRP--- 307
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 235 GFSGSDMDLHRMSRMSQKEYRPVPRPSYRQKVAAHNSPFSSEKALSTSSNLGSOYO-- 299
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 308 ---GKSLRLINKSLADSGEYMKYISKJGNDASANLTI--VENASTSTSTT--GTSHLVK 362
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY      293 -----CENG-----VSGPSQDLLPPNPY 311
          | | : | | :
Db      363 CAERKEKTCVNGGCEPMVKDLSNPSRYLCKCPNDF 397

```

```

RESULT 10
US-08-467-602-323
Sequence 323, Application US/08467602C
Patent No. 6444642
GENERAL INFORMATION:
APPLICANT: Sklar, Robert
APPLICANT: Marchionni, Mark
APPLICANT: Gwynne, David I.
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
FILE REFERENCE: 04585/028003
CURRENT APPLICATION NUMBER: US/08/467,602C
CURRENT FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/209,204
EARLIER FILING DATE: 1994-03-08
EARLIER APPLICATION NUMBER: 08/059,022
EARLIER FILING DATE: 1993-05-06
NUMBER OF SEQ. ID NOS. 420
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 323
LENGTH: 601
TYPE: PRP
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (601)...(601)
OTHER INFORMATION: xaa is Arg or absent.
US-08-467-602-323

```

Query Match	5.2%	Score 98.5	DB 4	Length 601
Best Local Similarity	23.0%	Pred. No. 0.07		
Matches 77	Conservative 40	Mismatches 119	Indels 99	Gaps 18

[illegible]

RESULT 11
US-08-470-335-236
Sequence 236, Application US/08470335E
Patent No. 6147190
GENERAL INFORMATION:
APPLICANT: GOODEARL, ANDREW
APPLICANT: STROCKBANT, PAUL
APPLICANT: MINGHETTI, LUISA
APPLICANT: WATERFIELD, MICHAEL

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:49:08 ; Search time 21 Seconds
(without alignments)
1776.837 Million cell updates/sec

Title: US-10-027-059A-1
Perfect score: 1889
Sequence: 1 MADADGFGIAHTPLEPDAK.....CDHPMSICPLSTYLSLGMG 349

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1889	100.0	349	9 US-10-027-059A-1	Sequence 1, Appl1
2	1767	93.5	518	9 US-10-027-059A-3	Sequence 3, Appl1
3	535	28.3	530	9 US-10-008-264A-4	Sequence 4, Appl1
4	528.5	28.0	535	9 US-10-008-264A-2	Sequence 2, Appl1
5	517.5	27.4	365	10 US-09-925-300-1676	Sequence 1676, App
6	158	8.4	57	10 US-09-864-761-33963	Sequence 33963, A
7	198	5.2	24	9 US-09-983-802-522	Sequence 522, App
8	98.5	5.2	418	9 US-09-946-807-3	Sequence 3, Appl1
9	98.5	5.2	418	10 US-09-795-686-3	Sequence 3, Appl1
10	98.5	5.0	418	10 US-09-795-686-3	Sequence 3, Appl1
11	94	5.0	30	10 US-09-864-761-39230	Sequence 39230, A
12	93	4.9	320	10 US-09-854-122-20	Sequence 20, Appl1
13	92.5	4.9	1252	9 US-09-908-193-19	Sequence 19, Appl1
14	92.5	4.8	1253	9 US-09-908-193-20	Sequence 20, Appl1
15	91.5	4.8	588	10 US-09-801-368-46	Sequence 46, Appl1
16	91.5	4.8	1252	9 US-09-908-193-21	Sequence 21, Appl1
17	91.5	4.8	1252	10 US-09-754-997A-2	Sequence 2, Appl1
18	89	4.7	1115	9 US-10-174-990-440	Sequence 440, App
19	89	4.7	1115	9 US-10-176-758-440	Sequence 440, App

20	89	4.7	1115	9 US-10-175-737-440	Sequence 440, App
21	89	4.7	1115	9 US-10-173-706-440	Sequence 440, App
22	89	4.7	1115	9 US-10-175-738-440	Sequence 440, App
23	89	4.7	1115	9 US-10-175-752-440	Sequence 440, App
24	89	4.7	1115	9 US-10-176-482-440	Sequence 440, App
25	89	4.7	1115	9 US-10-176-757-440	Sequence 440, App
26	89	4.7	1115	9 US-10-176-913-440	Sequence 440, App
27	89	4.7	1115	9 US-10-180-552-440	Sequence 440, App
28	89	4.7	1115	9 US-10-180-557-440	Sequence 440, App
29	89	4.7	1115	9 US-10-173-700-440	Sequence 440, App
30	89	4.7	1115	9 US-10-174-572-440	Sequence 440, App
31	89	4.7	1115	9 US-10-174-579-440	Sequence 440, App
32	89	4.7	1115	9 US-10-174-582-440	Sequence 440, App
33	89	4.7	1115	9 US-10-174-588-440	Sequence 440, App
34	89	4.7	1115	9 US-10-175-739-440	Sequence 440, App
35	89	4.7	1115	9 US-10-175-740-440	Sequence 440, App
36	89	4.7	1115	9 US-10-175-743-440	Sequence 440, App
37	89	4.7	1115	9 US-10-176-488-440	Sequence 440, App
38	89	4.7	1115	9 US-10-176-492-440	Sequence 440, App
39	89	4.7	1115	9 US-10-176-747-440	Sequence 440, App
40	89	4.7	1115	9 US-10-176-750-440	Sequence 440, App
41	89	4.7	1115	9 US-10-176-985-440	Sequence 440, App
42	89	4.7	1115	9 US-10-176-987-440	Sequence 440, App
43	89	4.7	1115	9 US-10-176-991-440	Sequence 440, App
44	89	4.7	1115	9 US-10-176-992-440	Sequence 440, App
45	89	4.7	1115	9 US-10-176-993-440	Sequence 440, App

ALIGNMENTS

RESULT 1
US-10-027-059A-1
Sequence 1, Application US/10027059A
Patent No. US20020165382A1
GENERAL INFORMATION:
APPLICANT: Basson, Craig
TITLE OF INVENTION: Transcription Factors that Regulate No. US20020165382A1 and
FILE REFERENCE: 955-12P
CURRENT APPLICATION NUMBER: US/10/027, 059A
CURRENT FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 349
TYPE: PRT
ORGANISM: Homo sapiens
US-10-027-059A-1

Query Match 100.0%; Score 1889; DB 9; Length 349;
Best Local Similarity 100.0%; Pred. No. 3 8e-168;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MADADGFGIAHTPLEPDAKDLPDCKSPESALGAPSKSPSSPOAFTQOGMEGIKVFLHE 60
|||||
DB 1 MADADGFGIAHTPLEPDAKDLPDCKSPESALGAPSKSPSSPOAFTQOGMEGIKVFLHE 60
QY 61 RELMLKFEHVEGTEMITTKAGRRMFPKYKVVYVTLNKRKTKIILMDIYPADDDHKYKADNK 120
|||||
DB 61 RELMLKFEHVEGTEMITTKAGRRMFPKYKVVYVTLNKRKTKIILMDIYPADDDHKYKADNK 120
QY 121 WSVTGAEPMPRLVYHPDPSPATGAHMMQVLSFOKLKLTNNHLDPEGHTILNSMHKYQ 180
|||||
DB 121 WSVTGAEPMPRLVYHPDPSPATGAHMMQVLSFOKLKLTNNHLDPEGHTILNSMHKYQ 180
QY 181 PRLHLYKADENNFGSKNTAFCTHVEPETAFIAVTSYONHKITQVLKLENNPFAKGRGSD 240
|||||
DB 181 PRLHLYKADENNFGSKNTAFCTHVEPETAFIAVTSYONHKITQVLKLENNPFAKGRGSD 240
QY 241 DMLHMRKRSQSKSEYVYVRSYTRQVYASNHSPSSBSRALSTSSNIGSYQGENYSGP 300
|||||
DB 241 DMLHMRKRSQSKSEYVYVRSYTRQVYASNHSPSSBSRALSTSSNIGSYQGENYSGP 300

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rapb

Page 2

```
QY 301 SODLLPPNNYPLPOEHSQIYHCTKRKGECDHPSICFLSYLPLSLGWC 349
Db 301 SODLLPPNNYPLPOEHSQIYHCTKRKGECDHPSICFLSYLPLSLGWC 349

RESULT 2
US-10-027-059A-3
; Sequence 3, Application US/10027059A
; Patent No. US20020165382A1
; GENERAL INFORMATION:
; APPLICANT: Basson, Craig
; TITLE OF INVENTION: Transcription Factors that Regulate No. US20020165382A1 and M
; FILE REFERENCE: 955-12P
; CURRENT APPLICATION NUMBER: US/10/027,059A
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 3
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-027-059A-3

Query Match 93.5%; Score 1767; DB 9; Length 518;
Best Local Similarity 97.9%; Pred. No. 1,7e-156;
Matches 331; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 MADADEGFLAHTPLEPDAKDLPCDSKPESALGAPSKSPSSPOAFTQOGMEGIVFLHE 60
Db 1 MADADEGFLAHTPLEPDAKDLPCDSKPESALGAPSKSPSSPOAFTQOGMEGIVFLHE 60

QY 61 RELWLKFEHVEGTEMITTKAGRRMPPSYKVYTGILNPKTKYLLMDIVPADDDRRYKFADNK 120
Db 61 RELWLKFEHVEGTEMITTKAGRRMPPSYKVYTGILNPKTKYLLMDIVPADDDRRYKFADNK 120

QY 121 WSVTGKAEPPAMPGRLYVHPDSPAATGAMHMRQVLSFOKLTNNHLDPPFGHIIILNSMHRKY 180
Db 121 WSVTGKAEPPAMPGRLYVHPDSPAATGAMHMRQVLSFOKLTNNHLDPPFGHIIILNSMHRKY 180

QY 181 PRLIIVKADENNNGSGSKNTAFCTHYEPETAFIATVSQNHKLTQKIENNPFAKGFGRSD 240
Db 181 PRLIIVKADENNNGSGSKNTAFCTHYEPETAFIATVSQNHKLTQKIENNPFAKGFGRSD 240

QY 241 DMELHRSRMSQSKYVYVPRSTYVROKVASNSPSSSSRALSTSSNLGSGYOCENGVSGR 300
Db 241 DMELHRSRMSQSKYVYVPRSTYVROKVASNSPSSSSRALSTSSNLGSGYOCENGVSGR 300

QY 301 SODLLPPNNYPLPOEHSQIYHCTKRKGE---CDHFW 334
Db 301 SODLLPPNNYPLPOEHSQIYHCTKRKGE---CDHFW 334

RESULT 3
US-10-008-264A-4
; Sequence 4, Application US/10008264A
; Publication No. US20030104528A1
; GENERAL INFORMATION:
; APPLICANT: Szabo, Laurie H.
; TITLE OF INVENTION: T-BET COMPOSITIONS AND METHODS OF USE THEREOF
; FILE REFERENCE: HUI-040CP
; CURRENT APPLICATION NUMBER: US/10/008,264A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: PCT/US00/15345
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137085
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 4
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-008-264A-4
```

```
US-10-008-264A-4
Query Match 28.3%; Score 535; DB 9; Length 530;
Best Local Similarity 35.2%; Pred. No. 1.9e-41;
Matches 147; Conservative 40; Mismatches 97; Indels 134; Gaps 16;

QY 16 EPDAKDLPCDSKPESALGAP-----SKSPSSPOAFT-----TQO 49
Db 38 EPDAKDLPCDSKPESALGAP-----SKSPSSPOAFT-----TQO 49

QY 50 GMEG-----IKFVLERELWLKFEHVEGTE 74
Db 97 GMEG-----IKFVLERELWLKFEHVEGTE 74

QY 134 RLYVHPDSPAATGAMHMRQVLSFOKLTNN---HDPFGHIIILNSMHRKYORPLHIVADE 190
Db 217 RLYVHPDSPAATGAMHMRQVLSFOKLTNN---HDPFGHIIILNSMHRKYORPLHIVADE 190

QY 277 -----GEPREACASNTVHFTFOETOFIAVAYONAEITQKIDNNPFAKGFRENFESMY 331
Db 277 -----GEPREACASNTVHFTFOETOFIAVAYONAEITQKIDNNPFAKGFRENFESMY 331

QY 238 -----GSDMELHRSRMSQSKYVY-----VPRSTYKOR--- 266
Db 332 ASVDTSPSPGPNCLLGDP-----FSPILSNQYVPSRFYRPLDLPQPRDMISQPYWL 386

QY 267 -VASNSPSSSSRALSTSSNLGSGYOCENGVSGRS-----QDLPNNYPLPOEHS 317
Db 387 GTREHS-YEAFRAVSMKPTL-----LPSAPGPTVYRYRODVLATGACMPVAPQY 437

RESULT 4
US-10-008-264A-2
; Sequence 2, Application US/10008264A
; Publication No. US20030104528A1
; GENERAL INFORMATION:
; APPLICANT: Szabo, Laurie H.
; TITLE OF INVENTION: T-BET COMPOSITIONS AND METHODS OF USE THEREOF
; FILE REFERENCE: HUI-040CP
; CURRENT APPLICATION NUMBER: US/10/008,264A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: PCT/US00/15345
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137085
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-008-264A-2

Query Match 28.0%; Score 528.5; DB 9; Length 535;
Best Local Similarity 37.7%; Pred. No. 7.8e-41;
Matches 140; Conservative 41; Mismatches 115; Indels 75; Gaps 12;

QY 3 DADE--GFLAHTPLEPDAKDLPCDSKPESALGAPSKSPSSPOAFTQOG-----ME 52
Db 44 DADE--GFLAHTPLEPDAKDLPCDSKPESALGAPSKSPSSPOAFTQOG-----ME 52

QY 53 G-----IKFVLERELWLKFEHVEGTEMIT 77
Db 101 G-----IKFVLERELWLKFEHVEGTEMIT 77

QY 78 KAGRRMFPSSYKVYVGNPKTKYLLMDIVPADDDRRYKFADNKSVTGKAEPPAMP--RLY 136
Db 161 KAGRRMFPSSYKVYVGNPKTKYLLMDIVPADDDRRYKFADNKSVTGKAEPPAMP--RLY 136
```

QY 137 VHPSPATGAMHMRQVLSFQKTLKLTNN---HLDPSTILNSMHRKYQRLHIVKADENN 193
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
DB 221 VHPSPNPGAMHMRQVLSFQKTLKLTNNKSGASNNYQMLYGLKHYQRLHIVKADENN 280
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 194 FGSKNTAFCTHV--FPEFAFIATVSYQNHKTOLKIENNPAKFGSGDDMLHMRMO 251
:
DB 281 EAACH-ASNTHTFFQETQFIATVAYQNAELTQKIDNNPFKGFREN---FEEMYSV 335
:
QY 252 SKKEYVVRSTYRQKVASNHSPSSSRALSTSSNLGSOYOCEN---GVSPSQDLLP 307
:
DB 336 DTSTSPGNCQFGLSGDHSPL-----LPNQYPPVSRFPDLPQADKDVVQ 383
:
QY 308 PNPYPLPOEHS 318
:
DB 384 AYWLGPARDHS 394
:

RESULT 5
US-09-925-300-1676
: Sequence 1676, Application US/09925300
: Patent No. US20020151681A1
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben,
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA101
: CURRENT APPLICATION NUMBER: US/09/925,300
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05988
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1890
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 1676
: LENGTH: 365
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (47)
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (220)
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1676

Query Match 27.4%; Score 517.5; DB 10; Length 365;
Best Local Similarity 45.9%; Pred. No. 4,7e-40;
Matches 112; Conservative 33; Mismatches 74; Indels 25; Gaps 4;

QY 83 MFPSTKVVYGLNPKTKYILLMDIYPADDRYKFDNKNKSVTGKAEPAWGLYVHPDSP 142
:
DB 4 MFPSTKVVYGLNPKTKYILLMDIYADDCRYKFNHSMVAGXADPEMKRMYIHPDSP 63
:

QY 143 ATGAMMQLVSEFKTLKLTNNHLDPEG-----HTLNSMHRKYQRL 182
:
DB 64 ATGEGMMSKVYTFHKLKLTNNISDKHGTFLAPPSDHATWQNYSFQTQTLNSMHRKYQRL 123
:

QY 183 LHIIVADENNFGSKNTAFCTHVPEFAFIATVSYQNHKTOLKIENNPAKFGSGDDM 242
:
DB 124 FHIIVR--NDILKLYSTFRITLYLPETETRIATAYQNKITQKIDNNPFKGFRTDNG 181
:

QY 243 ELHMRSMQSKETPYVVPSTYRQKVASNHSPSSSRALSTSSNLGSOYOCENGVSPSQ 302
:
DB 182 RREKHKQLTLOSRAVYFDERHKKENGTSDS--SSDOAFNXPQAQSSPASTVGTSN-LK 238
:

QY 303 DLLP 306
:
DB 239 DLCP 242
:

Query Match 8.4%; Score 158; DB 10; Length 57;
Best Local Similarity 45.9%; Pred. No. 4,7e-40;
Matches 112; Conservative 33; Mismatches 74; Indels 25; Gaps 4;

RESULT 6
US-09-864-761-33963
: Sequence 33963, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecmca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 33963
: LENGTH: 57
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC007076.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN BT74, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
: OTHER INFORMATION: EST_HUMAN HIT: BE468028.1, EVALUE 5.00e-17
: OTHER INFORMATION: SWISSPROT HIT: Q94890, EVALUE 2.00e-19
US-09-864-761-33963

Query Match 8.4%; Score 158; DB 10; Length 57;
Best Local Similarity 45.9%; Pred. No. 4,7e-40;
Matches 112; Conservative 33; Mismatches 74; Indels 25; Gaps 4;

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rapb

Page 4

Best Local Similarity 50.9%, Pred. No. 1.2e-07;
Matches 28; Conservative 14; Mismatches 11; Indels 2; Gaps 1;
QY 82 RMEPTKRVYGLNPKRTYLLMDIVADHRKFKADNK--WSYTGKKEPMPGR 134
DB 1 RMEPTLRVSFGVDEAKTIVLMDIVADHRKFKATVHSSMLVAGKADPELPAR 55
RESULT 7
US-09-983-802-522
Sequence 522, Application US/09983802
Publication No. US20030022185A1
GENERAL INFORMATION:
APPLICANT: Stiegher, et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE NO. OF INVENTION: 2001091
CURRENT FILING DATE: 2001-10-25
PCT/US98/13684
PCT FILING DATE: 1998-07-07
PCT APPLICATION NUMBER: 60/051,926
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,925
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,929
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/052,803
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/052,732
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,931
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,932
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,916
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,930
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,918
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,920
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/052,733
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/052,795
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,919
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/051,928
PRIORITY FILING DATE: 1997-07-08
PRIORITY APPLICATION NUMBER: 60/052,722
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,723
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,948
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,949
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,953
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,950
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,947
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,964
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/056,360
PRIORITY FILING DATE: 1997-08-18
PRIORITY APPLICATION NUMBER: 60/055,664

PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,785
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,664
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 522
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
US-09-983-802-522
Query Match
Best Local Similarity 5.2%, Score 99; DB 9; Length 24;
Matches 18; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 214 VSYONHKITOLKIENPFKGR 237
DB 1 VSYONQITRLKIDRNPFAKGR 24
RESULT 8
US-09-946-807-3
Sequence 3, Application US/09946807
Patient No. US2002016514A1
GENERAL INFORMATION:
APPLICANT: Stefansson, Hreinn
APPLICANT: Gulcher, Jeffrey R.
TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
FILE REFERENCE: 2345, 2004-001
CURRENT APPLICATION NUMBER: US/09/946,807
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US/09/795,668
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 09/515,716
PRIOR FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 1531
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 3
LENGTH: 418
TYPE: PRT
ORGANISM: Homo sapiens
US-09-946-807-3
Query Match
Best Local Similarity 5.2%, Score 98.5; DB 9; Length 418;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;
QY 25 DSIPES---LUGAPSKS-----PSSPOAFTOGMEGIVFLAEREILKHEVG 71
DB 110 DRPPAPAGRAIGPAEPLIANGTVPSWPTAPVPSAGEG-----EAPVLYKVHGV 163
QY 72 TEMITLKAGRMFYSYKVTYGLNPKRTYLLMDIVPADHRKFKADNKWSYTGKAEPM 131
DB 164 -----MAVKAGGL--KKDSLTVRL-----GTW-----GHPAF 189
QY 132 P--GR-----YVHPDSPATGAMHMQQLVSFOKRLTKNNHLDPRGHIILNSMKRKYOP 181
DB 190 PSCGRLEKEDSRVIFEMEDDANSTSRAPAFRASFPPL--TGRNLKKEVSVLCKRCALPP 248
QY 182 RLHIYVADENNFGSKNTACTVFPETAFIVTSYON-----HKITOLKIENPFK 234
DB 249 RLKEMKSOE--SAAGSKIVLC--ETSSYSSLRKFKWKNGNELNKRKNPOMIKIOKKP--- 303


```

: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 39230
: LENGTH: 30
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC005826.1
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
: OTHER INFORMATION: EST_HUMAN HIT: AM664610.1, EVALU6 7.00e-05
: OTHER INFORMATION: SWISSPROT HIT: Q9UMR3, EVALU6 1.00e-11
: US-09-864-761-39230
: Query Match
: Best Local Similarity 5.0%; Score 94; DB 10; Length 30;
: Matches 19; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
:
: OY 51 MEGIKVFEHERELMKFHEVETEMITKAG 80
: DB 1 MAKIACSLKTKLMKPFHELTETEMITKSC 30
:
: RESULT 12
: US-09-854-122-20
: Sequence 20, Application US/09854122
: Patent No. US20020016980A1
: GENERAL INFORMATION:
: APPLICANT: SMITH, ROBERT
: TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
: FILE REFERENCE: PHA-007.01
: CURRENT APPLICATION NUMBER: US/09/854,122
: PRIOR FILING DATE: 2001-09-10
: PRIOR APPLICATION NUMBER: 60/202,529
: PRIOR FILING DATE: 2000-05-10
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 20
: LENGTH: 320
: TYPE: PRT
: ORGANISM: Flaveria bidentis
: US-09-854-122-20
:
: Query Match
: Best Local Similarity 4.9%; Score 93; DB 10; Length 320;
: Matches 60; Conservative 28; Mismatches 97; Indels 74; Gaps 13;
:
: OY 114 YKFAADKMSVYTGKAPAPGR--LYVHDS-----PATGAHMMBOLVSTOKLKTNNHL 165
: DB 33 YKFDYD-FMGLGNMTEGAILAQOSFKARPDVFLCSPKSGTJWLKALA---YAIIVTRKRF 88
```

```

: OY 166 DFF-CHITLNSMRYOPRLH-----IVKAEDNNGFGSKNTAFCTHVEPETATF-----AVT 215
: DB 89 DEFTSPLLTNIPHCIPYIEKDLKIVENONNSCTPMATIMPYHVLPSILALNCKMY 148
:
: OY 216 SYONHK-----ITQKTIENNPPAKGFRGSDMELHRSRM-----Q 251
: DB 149 IYRNKIDVIYSEYHFGREITKLPLEDAPFEAAV---DEFYHGISOFGPYWHLGYNKA 204
:
: OY 252 SKREYPVV-----PRSTVRQKVASNHSPPSSSR-----ALSTSSNLGSO 290
: DB 205 SLREPEVILFLKYEDYKDKDPTSNVKRLAEFIQYPTFEEKEGVIESIILKCSFENLSN- 263
:
: OY 291 YOCENGVSQPSODLLPPPN 309
: DB 264 --LEVNRKSGNSKGFLPIEN 280
```

RESULT 13

US-09-908-193-19

Sequence 19, Application US/09908193
Publication No. US20020192748A1

GENERAL INFORMATION:

APPLICANT: RASTELLI, LUCA

APPLICANT: SHIRKETS, RICHARD A.

APPLICANT: ZERNHUSEN, BRIAN

APPLICANT: MALYANKAR, BRIEL M.

APPLICANT: PADIGARU, MORALIDHARA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

FILE REFERENCE: 21402-062

CURRENT APPLICATION NUMBER: US/09/908,193

PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: 60/220,273

PRIOR FILING DATE: 2000-07-24

PRIOR APPLICATION NUMBER: 60/221,650

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/221,233

PRIOR FILING DATE: 2000-07-27

PRIOR APPLICATION NUMBER: 60/220,912

PRIOR FILING DATE: 2000-07-26

PRIOR APPLICATION NUMBER: 60/218,875

PRIOR FILING DATE: 2000-07-18

PRIOR APPLICATION NUMBER: 60/218,870

PRIOR FILING DATE: 2000-07-18

PRIOR APPLICATION NUMBER: 60/218,901

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 19

LENGTH: 1252

TYPE: PRT

ORGANISM: Mus musculus

US-09-908-193-19

Query Match

Best Local Similarity 4.9%; Score 92.5; DB 9; Length 1252;

Matches 92; Conservative 44; Mismatches 126; Indels 183; Gaps 22;

```

: OY 11 AHPTLEPPAKDLPCDSKRESALGAPSKSPSOAAFTOOGEGIKVFLHRELMLEFHEV 70
: DB 626 SHVPFAP--AEUKYRAKMESLVSWQRPHPHQ-----ISGKLV-----WREV 667
:
: OY 71 GTEMIT---TKAGRMFPYSKYKVTGLNPKTKYLLMDIVPADDHRYKF--ADNK----- 120
: DB 668 GTEEDADGDRPPGGRGOANDVGPVRLKKKYQYELDLYVGRLEYKLVAFKHKHEDGYA 727
:
: OY 121 --WSVTGKAE-----PAMPGRLYVHPDSPATGAMMR--QLVSTQKLKLTN--- 162
: DB 728 AAWK--GKTERAPPLDLPIDRGPPLP-PAHVHAESNSSTIWLKMKRPDTTVKIVNTIV 784
:
: OY 163 -----NH---LD--PFGHITLNSMRY 179
: DB 785 RGPWGLRNASLVITYTSSGCDILLIGLKPFTKYEFVAVQSHGVMDGPFQSVVERSTLPD 844
```


Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rpb

Page 8

Db	463	LDSPPOS	PRRSS	SSIT	KKG	RRS	-----	SGSS	-----	PT	RR	TR	VC	VS	CH	SS	505
Oy	327	KGEC	DHP	-WS	335												
Db	506	DSPCM	RRPS	WS	515												

Search completed: June 17, 2003, 10:03:51
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:44:00 ; Search time 19 Seconds
(without alignments)
1765.838 Million cell updates/sec

Title: US-10-027-059A-1

Perfect score: 1889

Sequence: 1 MADADEGFLAHTPLEPDAK.....CDHPWSTCFLSYFLSLGWMG 349

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	958	50.7	184	2	S72231		transcription fact
2	865.5	45.8	173	2	S72230		transcription fact
3	658.5	34.9	423	2	B88450		protein F21H1.3 [
4	649.5	34.4	701	1	S46458		transcription fact
5	633.5	33.5	702	2	G01840		t-box protein 2 -
6	629.5	33.3	540	2	S72233		transcription fact
7	620	32.8	974	1	A40213		optic lobe develop
8	581	30.8	182	2	D56530		Ce-tbx-2 protein -
9	550.5	29.1	252	2	A88508		protein H1A12.4 [
10	549	29.1	344	2	A88023		protein T27A1.6 [i
11	544.5	28.8	436	1	S08156		gene t protein - m
12	541.5	28.7	432	1	A41056		brachyury homolog
13	536.5	28.4	423	1	A49125		gastrulation regul
14	536.5	28.4	681	2	I78558		hypothetical regul
15	486	25.7	720	1	A55160		Ttg protein - frui
16	474.5	25.1	664	2	T29011		hypothetical prote
17	419	22.2	182	2	C56530		Ce-tbx-7 protein -
18	391	20.7	479	1	S41015		transcription fact
19	384.5	20.4	342	2	C88579		protein Lbx-8 [imp
20	381.5	20.2	377	1	S41019		transcription fact
21	340	18.0	352	2	T31544		hypothetical prote
22	282.5	15.0	303	2	T19424		hypothetical prote
23	267.5	14.2	322	2	E88473		protein F40H6.4 [i
24	221.5	11.7	376	2	T15166		hypothetical prote
25	171.5	9.1	302	2	T31542		hypothetical prote
26	141.5	7.5	507	2	T26809		hypothetical prote
27	133.5	7.1	331	2	T28023		hypothetical prote
28	123	6.5	414	2	T33067		hypothetical prote
29	123	6.5	414	2	E89467		protein ZK380.1 [i

30 120 6.4 181 2 T27769 hypothetical prote
31 120 6.4 447 2 AH0536 conserved hypothe
32 119 6.3 233 2 T28914 hypothetical prote
33 115.5 6.1 747 2 T33488 hypothetical prote
34 109.5 5.8 1489 2 G71406 probable retroviro
35 102 5.4 1651 2 T14160 transmembrane rece
36 101.5 5.4 533 2 T50401 hypothetical coile
37 100.5 5.3 642 2 T39607 fork head protein
38 98.5 5.2 509 2 AE1295 Bifunctional phosph
39 98 5.2 1454 2 E84535 probable retrocelem
40 96.5 5.1 509 2 AC1667 Bifunctional phosph
41 95 5.0 694 2 S32958 Probable membrane
42 94.5 5.0 422 2 S32357 glial growth facto
43 94.5 5.0 698 2 A54796 regulatory protein
44 94.5 5.0 1139 1 S61918 protein kinase C (
45 94 5.0 272 2 T40644 hypothetical prote

ALIGNMENTS

RESULT 1

S72231

transcription factor tbx5 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Nov-1999

C:Accession: S72231

R:Agulnik, S.I.; Garvey, N.; Hancock, S.; Ruvinisky, I.; Chapman, D.L.; Agulnik, I.; B
Genetics 144, 249-254, 1996

A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A:Reference number: S72230; MUID:97032942; PMID:8878690

A:Accession: S72231

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-184 <AGU>

A:Cross-references: EMBL:U57330; NID:gl620599; PIDN:AAC53109.1; PID:gl620600

A:Note: residue 184 is not shown

C:Genetics:

A:Gene: tbx5

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-184/Domain: T-box homology (fragment) <TBX>

Query Match 50.7%; Score 958; DB 2; Length 184;

Best Local Similarity 96.7%; Pred. No. 1.9e-70;

Matches 178; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 63 LWLAFHEVGTETMITKAGRRMPPSYKVKVTGLNPKTKYILLMDIYPADDDHRYKFKADNKNWS 122

Db 1 LWLFRSGVTETMITKAGRRMPPSYKVKVTGLNPKTKYILLMDIYPADDDHRYKFKADNKNWS 60

QY 123 VTGAEPAMPGRLYVHPDSPATGAHMMQLVSKLTKLTNNHLDLPFGHLLNSMHKYOPR 182

Db 61 VTGAEPAMPGRLYVHPDSPATGAHMMQLVSKLTKLTNNHLDLPFGHLLNSMHKYOPR 120

QY 183 LHIVKADENNGSGKNTAFCTHVPETAFIAVTSYQNHKITOLKIENPNPFAKFGSGDDM 242

Db 121 LHIVKADENNGSGKNTAFCTHVPETAFIAVTSYQNHKITOLKIENPNPFAKFGSGDDL 180

QY 243 ELHR 246

Db 181 ELHR 184

RESULT 2

S72230

transcription factor tbx4 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Nov-1999

C:Accession: S72230

R:Agulnik, S.I.; Garvey, N.; Hancock, S.; Ruvinisky, I.; Chapman, D.L.; Agulnik, I.; B
Genetics 144, 249-254, 1996

A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A: Map position: 17q21-17q22
C: Superfamily: mouse transcription factor tbx2; T-box homology
F: 104-285/Domain: T-box homology <TBX>

Query Match 33.5%; Score 633.5; DB 2; Length 702;
Best Local Similarity 52.0%; Pred. No. 2.4e-43;
Matches 130; Conservative 30; Mismatches 53; Indels 37; Gaps 3;
Qy 23 PCDSKPESALGAPSKSP-----SSPQAFTQGM----- 51
Db 29 PALALPPGALAKPLPLDGLAGAAAAAAGLHVSALGPHPPAAHLRLSLKLEPE 88
Qy 52 ----EGIKVFLHERELMLKHEVGTETMLITKAGRMPPSYKVTGLNPKTKYLLMDIV 107
Db 89 DEVEDDPKVTLEAKELQDFHKLGTENVITKSGRMPPFKRVSGLDKAKYLLMDIV 148
Qy 108 PADDRHYKFNADNKSVTGKAPAMPGRGLYVHPDSPATGAHWMROLVSFQKLKLTNNHLD 167
Db 149 AANDCRKFNINSRWVAGKADPENPKRMYIHPDSPATGEQMAKPVAFHKLKLTNNISDK 208
Qy 168 FGHILNSMHKYQPRHLHVKADENNGFGSKNTAFCTHVFPEAFIATVSYONHKITOLKI 227
Db 209 HGFTILNSMHKYQPRHLHVRA--NDILKLPYSTERTVEPETDFIATVAYONDKITOLKI 266
Qy 228 ENNPFAGFR 237
Db 267 DNNPFAGFR 276

RESULT 6

S72233
transcription factor tbx6 - mouse
C: Species: Mus musculus (house mouse)
C: Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C: Accession: S72233; S72232
R: Agulnik, S.I.; Chapman, D.L.; Hancock, S.; Silver, L.M.
submitted to the EMBL Data Library, May 1996
A: Reference number: S72233
A: Accession: S72233
A: Molecule type: mRNA
A: Residues: 1-540 <AGU>
A: Cross-references: EMBL:U57331; NID:91620601; PIDN: AAC53110.1; PID: g1620602
R: Agulnik, S.I.; Garvey, N.; Hancock, S.; Kuvinsky, I.; Chapman, D.L.; Agulnik, I.; Boll
Genetics 144, 249-254, 1996
A: Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.
A: Reference number: S72230; MUID: 97032942; PMID: 8878690
A: Accession: S72232
A: Status: nucleic acid sequence not shown
A: Molecule type: mRNA
A: Residues: 100-280 <AGW>
A: Cross-references: EMBL:U57331
C: Genetics: tbx6
C: Superfamily: T-box homology
C: Keywords: DNA binding
F: 100-282/Domain: T-box homology <TBX>

Query Match 33.3%; Score 629.5; DB 2; Length 540;
Best Local Similarity 43.4%; Pred. No. 3.6e-43;
Matches 138; Conservative 45; Mismatches 80; Indels 55; Gaps 9;
Qy 11 AHTPLEDAKDLPCDSKPESALGAPSKSPSPQAFTQQGMEGKVFLEHERELMKFHEV 70
Db 62 AAAPLPSS-----PIALG-PETAPPPEAL---HSLPGVSLSENQELWKEFSV 107
Qy 71 GTEMLITKAGRMPPSYKVTGLNPKTKYLLMDIVPADDRHYKFNADNKSVTGKAEPA 130
Db 108 GTEMLITKAGRMPPACRVSVTGDPEARFLFLDVPVGDGARYRWGPDWEPESGKAEP 167
Qy 131 MPGRLYVHPDSPATGAHWMROLVSFQKLKLTNNHLDPFHILNSMHKYQPRHLHVRADE 190
Db 168 LPDRVYIHPDSPATGAHWMROLVPSFHRVKLTNNSTLDPHGLHILNSMHKYQPRHLHVRA-- 225

Qy 191 NNGFGSKNTAFCTH-----VEPETAFIATVSYONHKITOLKIENPNFAKGR- GSD 241
Db 226 -----TQLCSQHWGCVASFPEPTTIFISVAYQNPRITQLKIAANPFAKGRNGR 277
Qy 242 MELHRMSRMOSKEYPVVPRSTVROKVASNHSPPFSESRALSTSSNLGSOQCENGVSGPS 301
Db 278 CKRERDARVRRKLRGPEFVAT--EACGSGDTPGG-----PCDSTLGGDI 319
Qy 302 QDLPLPPNPYP-----LPQ 315
Db 320 RDSDEQAPTPGKLLLPQ 337

RESULT 7

A40213
optic lobe development omb protein - fruit fly (Drosophila melanogaster)
N: Alternate names: omb protein
C: Species: Drosophila melanogaster
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C: Accession: A40213; S34827
R: Pflugfelder, G.O.; Roth, H.; Poock, B.; Kerschner, S.; Schwarz, H.; Jonschker, B.; H
Proc. Natl. Acad. Sci. U.S.A. 89, 1199-1203, 1992
A: Title: The lethal(1)optomotor-blind gene of Drosophila melanogaster is a major orga
A: Reference number: S34827; MUID: 92159016; PMID: 1741374
A: Accession: A40213
A: Status: preliminary; not compared with conceptual translation
A: Molecule type: mRNA
A: Residues: 1-974 <PFL>
A: Cross-references: GB:M81796; NID: g158018; PIDN: AAA28736.1; PID: g158019
A: Note: sequence extracted from NCBI backbone (NCBIP: 82056)
R: Poock, B.; Ballew, J.; Pflugfelder, G.O.
Mol. Gen. Genet. 238, 325-332, 1993
A: Title: Transcript identification in the optomotor-blind locus of Drosophila melanog
A: Reference number: S34827; MUID: 93261414; PMID: 8492800
A: Accession: S34827
A: Molecule type: DNA
A: Residues: 1-447 <POE>
A: Cross-references: GB: S61732; NID: g402317; PIDN: AAB26697.1; PID: g402318
A: Experimental source: Larva
C: Genetics:
A: Gene: FlyBase: bi
A: Cross-references: FlyBase: FBgn0000179
C: Superfamily: optic lobe development omb protein; T-box homology
C: Keywords: DNA binding
F: 337-521/Domain: T-box homology <TBX>

Query Match 32.8%; Score 620; DB 1; Length 974;
Best Local Similarity 46.7%; Pred. No. 4.7e-42;
Matches 135; Conservative 38; Mismatches 70; Indels 46; Gaps 6;

Qy 55 KVFLEHERELMLKPFHEVGTETMLITKAGRMPPSYKVTGLNPKTKYLLMDIVPADDRHY 114
Db 329 KVTLEGRDLMEKFKHKLGTENVITKSGRMPPFQMKFRVSGLDKAKYILLDIVAADRY 388
Qy 115 KPADNKSVTGKAEPAAMPGRGLYVHPDSPATGAHWMROLVSFQKLKLTNNHLDPFH 171
Db 389 KFHNSRWVAGKADPENPKRMYIHPDSPTTGEQMAKPVSPHKLKLTNNISDKHGFVST 448
Qy 172 ILNSMHKYQPRHLHVKADENNGFGSKNTAFCTHVFPEAFIATVSYONHKITOLKIENPN 231
Db 449 ILNSMHKYQPRHLVRA--NDILKLPYSTERTVEPETDFIATVAYONDKITOLKIDNN 506
Qy 232 FAKGR--GSDMELHRMSRMOSKEYPVVPRSTVROKVASNHSPPFSESRALSTSSNLGS 289
Db 507 FAKGLRDTGACKRE-----KNCYRQALMSNR---GSDSKLNTHTVSSS 547
Qy 290 QYQENGVSGPSQDLLP-----PPNPYPLPQHS 318
Db 548 RAPLHLGAGRPPLHHPHAALLDNOQDDDKLLDVGVPQSPPLPLSHS 596

RESULT 8

D56530

Query Match 28.8%; Score 544.5; DB 1; Length 436;
Best Local Similarity 43.8%; Pred. No. 2.1e-36;
Matches 117; Conservative 39; Mismatches 88; Indels 23; Gaps 5;

QY 54 IKVFLHERELWLKPFHEVGTETMIITKAGRRMPPSYKVTGLNPKTKYILLMDIVPADDDR 113
DB 42 LRVGLESELRLFKELTNEMIVTKNGRRMFPVLKVNYSGLDPPNAMYSLFLDFVTADNHR 101
QY 114 YKFDNKKSVTGKAPAMPGRILYVHPDSPATGAHMMRLVFSQKLKLTNNHLDPEGHIL 173
DB 102 WKYNGEWPVGGKPEQAPSCVYIHPDSPNFGAHHMKAPVSEKVKLINK-LNGGGQIML 160
QY 174 NSMHYQPLRHIVKADENNGFGSKNTAFCTHVFETAFIATVSYONKHITQKLIENPPA 233
DB 161 NSLHKYEPRIHIVR-----GGPQMITSHCFEPETQFIATVAYQNEEITALKIKYNPFA 214
QY 234 KGFSGDDMLHR--MSRMOSKEYP-----VPRSTVRQKVASNHSPPFSESALSTS 284
DB 215 RAFLDAKERNNDHKOVNEEPGDCQOQYSGWGLVPGAGTLCPPASSHPQFG-----S 267
QY 285 SNLGSQYOCENGVSQSDLLPPNPY 311
DB 268 LSLPSTHCERYPALNRHRSPPSPY 294

RESULT 12
A41056
brachyury homolog - African clawed frog
N:Alternate names: gene T homolog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
R:Smith, J.C.; Price, B.M.J.; Green, J.B.A.; Weigel, D.; Herrmann, B.G.
Cell 67, 79-87, 1991
A:Title: Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response
A:Reference number: A41056; MUID:92005698; PMID:1711160
A:Accession: A41056
A:Molecule type: mRNA
A:Residues: 1-432 <SMI>
A:Cross-references: GB:M77243; NID:q213999; PIDN:AAA49663.1; PID:q214000
C:Comment: Both basic fibroblast growth factor (bFGF) and activin A induce expression of
C:Genetics:
A:Gene: Xbra
C:Superfamily: mouse gene T protein; T-box homolog
C:Keywords: DNA binding; mesoderm; transcription regulation
F:49-225/Domain: T-box homolog <TBX>

Query Match 28.7%; Score 541.5; DB 1; Length 432;
Best Local Similarity 41.2%; Pred. No. 3.6e-36;
Matches 113; Conservative 43; Mismatches 81; Indels 37; Gaps 5;

QY 54 IKVFLHERELWLKPFHEVGTETMIITKAGRRMPPSYKVTGLNPKTKYILLMDIVPADDDR 113
DB 40 LKVSLEERDLWTRFKELTNEMIVTKNGRRMFPVLKVNYSGLDPPNAMYSLFLDFVTADNHR 99
QY 114 YKFDNKKSVTGKAPAMPGRILYVHPDSPATGAHMMRLVFSQKLKLTNNHLDPEGHIL 173
DB 100 WKYNGEWPVGGKPEQAPSCVYIHPDSPNFGAHHMKAPVSEKVKLINK-MNGGGQIML 158
QY 174 NSMHYQPLRHIVKADENNGFGSKNTAFCTHVFETAFIATVSYONKHITQKLIENPPA 233
DB 159 NSLHKYEPRIHIVR-----GGTQMITSHCFEPETQFIATVAYQNEEITALKIKHNPPA 212
QY 234 KGFSGDDMLHRMSRMOSKEYPVPRSTVRQKVASNHSPPFSESALSTS----- 285
DB 213 KAF-----LDAKERNDY-----KOILDGIDQSNSFSQLTGLIPNGSGISCSPN 258
QY 286 -----NLGSQYOCENGVSQSDLLPPNPY 311
DB 259 HTQFGAPLSLSPHCERYSLNRHRSAPSPY 292

RESULT 13
A49125
gastrulation regulatory protein Zf-T - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A49125
R:Schulte-Merker, S.; Ho, R.K.; Herrmann, B.G.; Nusslein-Volhard, C.
Development 116, 1021-1032, 1992
A:Title: The protein product of the zebrafish homologue of the mouse T gene is expressed
A:Reference number: A49125; MUID:93201989; PMID:1295726
A:Accession: A49125
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423 <SC>
A:Cross-references: GB:S57147; NID:q299316; PIDN:AAB25829.1; PID:q299317
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:127982, NCBIP:127985)
C:Superfamily: mouse gene T protein; T-box homolog
F:44-220/Domain: T-box homolog <TBX>

Query Match 28.4%; Score 536.5; DB 1; Length 423;
Best Local Similarity 41.5%; Pred. No. 9e-36;
Matches 125; Conservative 41; Mismatches 90; Indels 45; Gaps 10;

QY 40 SSPQAFTQQGWE-----GIKVLHERELWLKPFHEVGTETMIITKAGRRMPPSYKVKVT 92
DB 15 SAVESEF-QKGSEKGDASERDIKLSLEDAELWTKFKELTNEMIVTKTGRMPFVLRSV 73
QY 93 GLNPKTKYILLMDIVPADDRHYKFKADNKKSVTGKAEPAKPGELYYVHPDSPATGAHMMRL 152
DB 74 GLDPNAMYSLVLLDFVAADNRRKVVYNGEWPVGGKPEQSCVYIHPDSPNFGAHHMKAP 133
QY 153 VSFQKLKLTNNHLDPEGHILNSMHKYQPLRHIVKADENNGFGSKNTAFCTHVFETAFI 212
DB 134 VSFQKVLKLNK-LNGGGQIMLSLHKYEPRIHIVK-----GGTQKMLSSQSPFETQFI 186
QY 213 AVTSYONKHITQKLIENPPAKFGFGSDDMELHRMSRMOSKEYPVPRSTVRQKVASNHS 272
DB 187 AVTAYQNEEITALKIKHNPPAKAF-----LDAKERSDHEKVP-----DHS 226
QY 273 PFSSSRALSTSSNLGSQYOCENGVSQSDLLPPN-NPYPLPQEHSQLYHCTKRKGEC 331
DB 227 TDNQO-----SGYSQLGGWFLPSNGMPGSPSS--PPQFNGAPV---HSSGYSERYSSURN 277
QY 332 H 332
DB 278 H 278

RESULT 14
I78558
hypoetical Brachyury (T)-related transcription factor - mouse
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C:Accession: I78558
R:Bulfone, A.; Smiga, S.M.; Shimamura, K.; Peterson, A.; Puellies, L.; Rubenstein, J.L.
Neuron 15, 63-78, 1995
A:Title: T-brain-1: a homolog of Brachyury whose expression defines molecularly distinct
A:Reference number: I58171; MUID:95344783; PMID:7619531
A:Accession: I78558
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-681 <RES>
A:Cross-references: GB:S78858; NID:q1222544; PIDN:AAA92011.1; PID:q1222545
C:Genetics:
A:Gene: Tbr-1/Tes-56
C:Superfamily: T-box homolog
F:213-401/Domain: T-box homolog <TBX>

Query Match 28.4%; Score 536.5; DB 2; Length 681;
Best Local Similarity 38.7%; Pred. No. 1.7e-35;
Matches 136; Conservative 53; Mismatches 105; Indels 57; Gaps 13;

Tue Jun 17 10:11:35 2003

```
QY 10 LAHTELEPDA--KDLPCDSKPESALCAPSKSPSSPQ-----AAF-----TOOGMEGI 54
Db 146 MAHPVITNGVNSLLNSPQ---GYPTAGYFYQOYGHYSYQCAPFYQFSSTQPLGLVPG 202
QY 55 K--VFLHERELWLKEHEVGTMIITKAGRMFSPSYKVKYGLNPKTKYILLMDIVPADDDH 112
Db 203 KAQVYLCNRLPLKLRHQTETMIITKQGREMFPLSFNLSGLDPTAHYNIFVDVILADPN 262
QY 113 RYKADNKSVTGKAEPAMPG-RLYVHPDSPATGAHWMROLVSFQKCLKLTNNHLDPEGH- 170
Db 263 HWRFGCKVPCGKADTVQGNRVYMHDPSPNTGAHWMRQEIFSGKCLKLTNNKGASNNNG 322
QY 171 --JILNSMHKYQRLHIYKADENNGFGSKNTA----FCTHVFPETAFIATVSYONHKITQ 224
Db 323 QMVVLSLHKYQRLHVVVEVNE--GTEDTSPQGRVQTFPTPETQFIATVAYQNTDITQ 379
QY 225 LKIENNPFAKGRFGS-----DMELHRMSRQSKYPPVP-----RSTVROKVASN 270
Db 380 LKIDHNPFAKGRDNYDTIYTGCDMDRLTPSPNDSPRSOIVPGARYAMAGSFLQDOFVSN 439
QY 271 HS-----PFSESRLSTSSNLGSOYOCEN-GVSGPSODLLPPPN 309
Db 440 YAKAREHPGAGPGGTDRSVPTHTNGLLSPQAEDEPGAPSPORWFTPAN 490

RESULT 15
A55160
Trig protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55160
R:Kispiert, A.; Herrmann, B.G.; Leptin, M.; Reuter, R.
Genes Dev. 8, 2137-2150, 1994
A:Title: Homologs of the mouse Brachyury gene are involved in the specification of poste
A:Reference number: A55160; MUID:95047361; PMID:7958884
A:Accession: A55160
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-720 <KIS>
C:Genetics:
A:Gene: FlyBase:Trig
A:Cross-references: FlyBase:Fbgn0011723
C:Superfamily: Drosophila Trig protein; T-box homology
F:96-272/Domain: T-box homology <TBX>

Query Match 25.7%; Score 486; DB 1; Length 720;
Best Local Similarity 35.2%; Pred. No. 2.3e-31;
Matches 102; Conservative 50; Mismatches 82; Indels 56; Gaps 6;

QY 54 IKVFLHERELWLKEHEVGTMIITKAGRMFSPSYKVKYGLNPKTKYILLMDIVPADDDH 113
Db 87 LRISLDDRELWLRFQNLTMETIVTKNGRMFPVVVKISASGLDPAAMYTVLLEFVIDSHR 146
QY 114 YKFADNKSVTGKAEPAMPGRRLYVHPDSPATGAHWMROLVSFQKCLKLTNNHLDPEGHIL 173
Db 147 WKYVNGEWPVGKAEVPPSNPIYVHPESPNEFGAHHMKPEISFAKVLTKN-TNGNGQMML 205
QY 174 NSMHKYQRLHIYKADENNGFGSKNTAFCTHVFPETAFIATVSYONHKITQKIENNPFA 233
Db 206 NSLHKYEPRLVLRV-----GSEORHVVTYPFPETOFIATVAYQNEEVTSLKIKYNPFA 259
QY 234 KGRGSDDMELHRMSRQSKYEP--VPRSTVROKVASNHSPPSESRLSTSSMLGSQY 291
Db 260 KAF-----LDKERPDLYPHDTHYGLWLPPTHTYAAAAA--- 298
QY 292 QCENGVSGPSODLLPPNPYPLPQHSQIYHC-----TKRGECD 331
Db 299 -----PPPLSTAQSHGLVASCPSVSSAGSVGPSGGSCD 332
```

Search completed: June 17, 2003, 09:49:27
Job time : 22 secs

LOCUS (LOC): HSU80987 GenBank (R)
GenBank ACC. NO. (GBN): U80987
GenBank VERSION (VER): U80987.1 GI:2281318
CAS REGISTRY NO. (RN): 186848-85-1
SEQUENCE LENGTH (SQL): 1050
MOLECULE TYPE (CI): mRNA; linear
DIVISION CODE (CI): Primates
DATE (DATE): 26 Jul 1997
DEFINITION (DEF): Human transcription factor TBX5 mRNA; complete cds.
SOURCE: human.
ORGANISM (ORGN): Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini;
Hominidae; Homo
NUCLEIC ACID COUNT (NA): 278 a 301 c 248 g 223 t
REFERENCE: 1 (bases 1 to 1050)
AUTHOR (AU): Basson, C.T.; Bachinsky, D.R.; Lin, R.C.; Levi, T.;
Elkins, J.A.; Soultz, J.; Grayzel, D.; Kroumpouzou, E.;
Traill, T.A.; LeBlanc-Straceski, J.; Renault, B.;
Kucherlapati, R.; Seidman, J.G.; Seidman, C.E.
TITLE (TI): Mutations in human TBX5 [corrected] cause limb and
cardiac malformation in Holt-Oram syndrome
JOURNAL (SO): Nat. Genet., 15 (1), 30-35 (1997)

THIS PAGE IS BLANK (USPTO)